In celebration of early middle English 'H'

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Neuphilologische Mitteilungen
IN CELEBRATION OF EARLY MIDDLE ENGLISH ‘h’

Abstract
In Old English the littera ‘h’ represented the reflexes of Germanic *x. The standard consensus is that these were foot initial [h] and post-nuclear [x~ç]. These functions were considerably expanded in early Middle English.

Since English at this period was not ‘classical’ or institutionalised it did not require fixed spelling: neither natural variation nor scribal eccentricity was prevented from surfacing in written forms. Early Middle English texts therefore display a wide range of representational strategies.

Aside from its retained historical uses, ‘h’ in early Middle English had a number of other functions. Certain scribes used ‘h’ for Ø; others used it as part of complex systems of diacritic representation in which it often appears to indicate either fricativeness in general or periods of ‘unplaced’ voicelessness within a syllable nucleus. In this paper we explore some of the segmental, diacritic and arguably ‘decorative’ uses of ‘h’.

1. Introduction
Early Middle English writing is often extremely complex and difficult to interpret. The complexity arises as a result of the post-Conquest hiatus in the writing of English. Scribes had to design new orthographies to represent the results of a century of massive and transformative phonological change. Since early Middle English had no supra-local standard it did not require fixed spelling: so different styles of orthographic design were able to flourish. And because unlike most modern standards one word / one spelling was not the norm, neither natural variation nor scribal eccentricity was prevented from surfacing in written forms. This allowed diverse spellings for a large proportion of the lexicon.

2. The weakness of [h]
In this paper we consider a littera whose phonetic properties are conducive to its developing multiple functions, and we illustrate the resulting complexity of its

---

1 These observations arise from work on early Middle English manuscript texts undertaken at the Institute for Historical Dialectology, Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh towards the compilation of A Linguistic Atlas of Early Middle English (LAEME). This research project was supported from 2000-006 by AHRC for which gratitude is here expressed. RL thanks the University of Cape Town for generous travel support. An earlier version of this paper was delivered at The 15th International Conference on English Historical Linguistics at Munich, August 2008.

2 We use the terminology of the medieval theory of littera (see Benskin 1997:1 note 1 and 2001:194 note 4). Our conventions for this paper are as follows. Single inverted commas enclose litterae; littera is the abstract or superordinate notion of the letter. Phonetic brackets enclose potestates, which are sound values, represented here by IPA symbols. Glosses and names of lexical categories are in small capitals. Etymological categories and citations are in italics.
uses in some early Middle English scribal systems. The phone [h] is commonly considered to be ‘weak’ for two reasons. First, it is the most likely consonant in an inventory to delete; secondly, it is rare for *h to be reconstructed for a proto-language. Most attested [h] are the lenitions of other, non-laryngeal segments (Lass 1976: chapter 6). In Old English the littera ‘h’ primarily represented the reflexes of Germanic *x. The consensus is that these were foot initial [h] and post-nuclear [x~ç]. As we will see it had some other functions as well, which were considerably expanded in early Middle English.

2.1 Initial ‘h’ loss and excrescent ‘h’
In some Middle English writing systems *h- often fails to appear in initial position where it is historically expected; contrariwise, ‘h’ often appears in historically vowel-initial words where it has no etymological source. These phenomena are two aspects of the same process: loss of initial [h] and consequent ‘hypercorrect’ employment of the now ‘non-referential’ littera ‘h’ in positions where it is not expected to be associated with a potestas. Thus (from Oxford, Bodleian Library, Digby 86): abben to have, heuele evil. Whether there were ever hypercorrect pronunciations with [h], as happens in some modern English dialects, is unknowable.

The identification of instances of loss and insertion is not, however, entirely unproblematic. Before assigning lexemes to categories, we have to decide what words qualify as having historical initial [h]. For native words there is normally only one possible etymology with respect to h-. But Middle English vocabulary also includes Anglo-French and Anglo-Latin borrowings (some ultimately from Greek). In early Middle English, French is the commonest source of Romance and Greek vocabulary; but often it is impossible to tell whether a word was borrowed via French or directly from Anglo-Latin. For our purposes, however, the proximate origin of these loans is not important: we cannot safely use Romance examples to illustrate the status of initial [h] in early Middle English. There was widespread loss of initial [h] in post-classical Latin and in all varieties of early Old French (Pope 1934: §§185, 730, 1237). Many Romance borrowings into early Middle English, therefore, had at least variant spellings without ‘h’; moreover, any surviving initial ‘h’ will not reliably imply phonetic [h]. In other words, there is no way of counting ‘h’-less spellings as showing deletion or ‘h’-full spellings as showing insertion. For this reason, our analysis must perforce be restricted to Germanic vocabulary.

2.2 The LAEME corpus materials
Of the 167 tagged texts in the LAEME corpus, 3 124 display one or the other or both of deleted or excrescent ‘h’ in native vocabulary. Of the 43 texts that show no sign of these processes, some are extremely short, which may account for lack

---

3 A total of nearly 640,000 tagged words in an overall word count of over 800,000.
of attestation. Some, however, are of considerable size; neither the main scribe of Bodley 34 (the B of AB language) nor the Titus text of the Ancrene Riwle show any examples of either deletion or insertion in the samples tagged for LAEME (12,833 and 14,085 tagged words respectively). So for this proportion of our texts, either [h]-loss has not happened, or historical spellings are 100% reliable in spite of it having happened.

In some texts the numbers of occurrences are extremely low. In the work of the Worcester Tremulous Scribe, the Worcester Fragments and Ælfric’s Grammar and Gloss (3,379 and 15,881 tagged words respectively) show one example of deletion each and none of insertion. All the above texts belong to the W Midlands. From the other side of the country, the Norfolk text The Bestiary (4,096 tagged words) has only one token of each process in lexical examples and only two dropped ‘h’ in personal pronouns. For those texts with very small counts, it is possible that a process of lexical diffusion may be just beginning.

Numbers and proportions of occurrences vary considerably in the 110 texts that show either or both processes. To form an accurate picture of the significance (if any) of such variation it would be necessary, for any one text, to count not just examples of insertion and deletion, and their type/token ratios, but also, for comparison, the number of potential contexts, and the numbers of historical spellings. It might be of interest to record the data from lexical and grammatical words separately. One could also divide the ‘h’ examples into those where the ‘h’ has been written as a part of the word in the ordinary manner, and those where it has been interlined as a scribal ‘correction’ (whether or not in words with historical ‘h’). One would also have to take account of deletion of initial ‘h’ (again whether or not in words with historical ‘h’). Outputs would then have to be normalised for text

4 One text that has already been given informal consideration is the East Midland Genesis and Exodus (Cambridge, Corpus Christi College 444). Milroy (1983:41-42) presents a ‘partial list’ of ‘h’-loss and ‘h’-insertion for this text and implies that the numbers of each are roughly equal. From the tagged text of fols. 1r-41r (the full text runs to 81r) we can confirm that in Genesis and Exodus the proportions are evenly balanced. Another substantial text that shows comparable and roughly equal numbers of ‘h’-loss and ‘h’-insertion is the work of Hand B of the Trinity Homilies (Cambridge, Trinity College B.14.52) also from the East Midlands. Some texts show considerable numerical imbalance between the two processes. Most commonly there are many more insertions than there are deletions, e.g. in the work of the South-West Midland scribe of Oxford, Bodleian Library, Digby 86 (22,292 tagged words) there are 107 types which show ‘h’ insertion and only 11 types that show ‘h’ deletion (cf. Laæamon B, London, British Library Cotton Otho C xiii, (placed in Wilts) where the LAEME sample (13,052 tagged words) shows ‘h’ deletion in 13 types and ‘h’ insertion in 31 types).

5 I.e. all words with historical [h] (for potential deletion) and all historically vowel initial words (for potential insertion).

6 Scribal ‘normalisation’ to a historical spelling denotes either a non-[h]-dropper correcting a non-historical exemplar spelling, or an [h]-dropper changing to a traditional spelling in spite of his own spoken usage. For Orm, who was the original writer of his verse homilies, the latter must be the case in the very few words to which he adds initial ‘h’ where he first omitted it (for evidence, not in the LAEME sample, see Johannesson 2000). In language 1 of the Lambeth
length. An investigation with this amount of detail would undoubtedly be possible using the LAEME corpus.

3. Systemic deployment of ‘h’

3.1 Weakness > loss/instability; instability > multiple uses

It is clear from the above that an early Middle English ‘h’ in initial position may represent historical [h] or Ø. A littera that loses its potestas may disappear, may be retained in its original orthographic (but no longer phonetic) function or may be redeployed. These behaviours can be either variable or categorical. Retention of initial ‘h’ may imply either of the following:

a. if the loss is variable, ‘h’ may continue (variably) to represent its historical value;

b. ‘h’ is redundant and (along with Ø) categorically represents Ø.

3.2 Diacritic uses of ‘h’

The (variable) loss of initial [h] frees the littera ‘h’ for redeployment, and early Middle English texts display a wide variety of adaptations. Some are already present in Old English, which took them over from Latin. ‘h’ was used diacritically in Latin, for the representation of the Greek aspirated stops: ‘ch’ for ‘χ’, ‘ph’ for ‘φ’ and ‘th’ for ‘θ’. When the aspirates spirantised in late antique times, diacritic ‘h’ in these contexts became associated with the new fricative realisations. Old English scribes, trained in writing Latin, could transfer these digraphs into the vernacular with the familiar phonetic denotations [x], [f] and [θ], and this practice continued into early Middle English: e.g. from Worcester Cathedral, Chapter Library Q 29, fols. 130v-131r, þech though, prophete prophet, heteth eat pres. ind. pl.

Homilies (South-West Midlands), there are only six instances of omitted initial ‘h’. Of these, two have been corrected by the insertion of ‘h’. A further example is shown where the word halie holy is immediately preceded on the previous line by ali subpuncted for deletion. This scribe also shows uncorrected excrescent initial ‘h’ in 11 types. Excrescent initial ‘h’ has been subpuncted for deletion in eorðe earth. In the Nero Wooing Group (Hand B of London, British Library, Nero A. xiv, also South-West Midlands) there are seven missing initial ‘h’ of which four have been corrected. Uncorrected excrescent ‘h’ appears in four types. It has been erased once before am AM (1st pers pres of be). With these texts, it is impossible to be sure whether the normalisations are corrections of exemplar forms or of the scribes’ own, but the uncorrected examples suggest at least variation in the scribes’ own dialects. The evidence here, and that in note 4 above, indicates that [h]-dropping was present across the country in early Middle English, though there is a tendency (Milroy 1983, 1992:198-199, followed by Johannesson 2000) to see it as a primarily eastern change.

Retention might also allow use of ‘h’ not phonetically but logographically, e.g. to distinguish homophones. But without direct knowledge of the phonetic situation in a particular text, such a purpose would not be transparent to us. It has been claimed (Crisma 2007) that such knowledge is in fact obtainable in certain syntactic environments. Crisma presents a study in which the presence or absence of initial [h] can be determined by the allomorphy of preceding indefinite articles and possessive pronouns.
Although ‘ch’ continued to be a possible representation of [x] in some scribal systems in early Middle English, it began to be employed much more widely for [tʃ].\(^8\) This usage was adopted from Anglo-French, and in Middle English is used in both borrowed and native words: e.g. chambre CHAMBER, cheken CHEEKS (Cambridge, Corpus Christi College 402: Ancrene Wisse). Probably by analogy with this, ‘sch’ or ‘sh’ (and more rarely ‘ssh’) begin to be used for [S], e.g. shome SHAME (Oxford Bodleian Library, Digby 86), schakeles SHACKLES (Corpus Ancrene Wisse), bisshup BISHOP (Cambridge, Trinity College B.14.52: Trinity Homilies, Hand B). Such spellings with added ‘h’ are found in addition to continued use of OE ‘sc’ and less common variants ‘ss’, and ‘s’, e.g. scritte SHRIFT, sisses FISH, sadue SHADOW (all from London, British Library, Arundel 292, The Bestiary). These uses of ‘h’ to indicate either a continuant or a partially continuant segment result in its being associated with fricativeness generally, the necessary specification of articulatory position being provided by the first element of the digraph.

Since ‘h’ was used to represent the reflexes of Gmc *x, it also appeared in the reflexes of *xl, *xn, *xr, *xw. Authorities differ on whether, by historical times, these were clusters or were merely voiceless sonorants. Since ‘h’ is associated not only with fricativeness but also with voicelessness, either interpretation is defensible. These items simplified or voiced before or during early Middle English. As part of their development, the graphic sequences representing them could reverse. In the LAEME corpus there are no examples of ‘hn’- or ‘nh’-. Instances of ‘hr’- are confined to two texts copied from Old English precursors: the Wintney Benedictine Rule (London, British Library, Cotton Claudius D iii) and the Worcester Tremulous Scribe’s version of Ælfric’s Grammar and Glossary (Worcester Cathedral, Chapter Library F 174, fol. 1r-63r). There are no examples of ‘rh’-spellings. ‘hl’- survives for rather longer; in LAEME 12 texts show ‘hl’-spellings, e.g. hlisteð LISTENS (< OE hlystedð) in Trinity Homilies, Hand B (late 12th century), and hlauerd LORD (< OE hlæford) in London, British Library, Egerton 613, Hand G, fol. 64r-70v (e): Poema Morale (first half of the 13th century). In 8 texts, the spelling ‘lh’- appears (three texts show both ‘hl’- and ‘lh’-spellings), e.g. lhoup LOWS (< OE hlōwð) in London, British Library, Harley 978, fol. 11v: Svmer is icumen in (mid 13th century) and lheȝȝe LAUGH (< OE hlæhhan) from London, British Library, Arundel 57: Ayenbite of Inwyth. This last text was written in 1340 in Canterbury, but it represents Kentish English of prior to 1300, the author Dan Michel being about 70 years old at the time of writing. Ayenbite is the latest example in the LAEME corpus to show such spellings.\(^9\)

---


\(^9\) There are only four later examples cited in MED, two in personal names, one s.v. lēne LEAN and the other s.v. lōf LOAF.
The reflex of OE hw- continues into early Middle English spelled at first with ‘h’ before ‘p/w’ (or rare ‘u’) and later increasingly with the ‘h’ following ‘p/w’. See e.g. the following spellings of which from the LAEME corpus: hpucche (London, British Library, Add 27909), hwich (Oxford, Jesus College 29), hyvch (Avenbiete of Inwyf), philche (Trinity Homilies, Hand A), whulc (London, British Library, Cotton Caligula A.ix, Lažamon A, hand A).\(^{10}\)

4. ‘h’ in Gonville and Caius Ancrene Riwle (G)

The G scribe of Ancrene Riwle has a ‘prodigal’ writing system, which abounds in variant spellings and complex sound/symbol relationships.\(^{11}\) Prodigal writing systems have tended in the past to be dismissed with one of two typical reactions. If the text in question is a copy of a work known in other versions in less complex systems, then the scribe’s multiple and/or eccentric spellings are a ‘chaotic’ failure to achieve what must have been his real intention. If there is no other less eccentric version for comparison, difficulties in interpreting the text may lead to a second reaction: that the scribe was ‘confused’ and was writing nonsense. The modern reader’s confusion must entail confusion in the writer. But complexity of system is not the same as lack of system; multivocal arrays frequently have discoverable internal structure. In Laing and Lass (2009: §§2.2.1, 4.3, 5.3), we analyse in detail the literal substitutions that occur in the G scribe’s text. For ‘h’ it shows all the historical and diacritic uses discussed above, as well as many less conventional ones. At first sight this seems an extraordinary example of prodigality – ‘h’ appears in 25 ‘different’ historical contexts within the G scribe’s complex system:

1. ‘h’ for historical initial [h], e.g. his HIS, heuen HEAVEN;
2. ‘h’ for historical medial and final [x], e.g. bipoht BETHOUGHT, pah THOUGH;
3. ‘h’ for historical medial and final [ç], e.g. mihte MIGHT, heh HIGH;
4. ‘h’ for historical medial [γ], e.g. dahes DAYS, -fuhel FOWL;
5. ‘h’ for historical medial [j], e.g. wiheles (< OE wigel) WILES, unphrihen UNWOUND;
6. ‘h’ in combination with ‘i’ for historical medial [j], e.g. geihet (< OE *gegan) CRY OUT (this example only beside gei-, gey- with no ‘h’);

\(^{10}\) Spellings of OE hw- words with initial ‘p/w’ without ‘h’ are also very widespread in early Middle English, e.g. pilk (London, British Library, Arundel 292: Three Sorrowful Things), wuch (Cambridge, Corpus Christi College 145, South English Legendary). ‘h’-less spellings with ‘qw/qu’ also occur locally, and with initial ‘u/v’ occasionally, e.g. Quilk (Cambridge, Corpus Christi College 444, Genesis and Exodus), vilche (Worcester Cathedral, Chapter Library Q 29).

\(^{11}\) For an introduction to the concepts economical and prodigal writing systems see LAEME Introduction, chapter 2, §2.3.2 and refs.
7. ‘h’ preceding ‘p’ for historical initial [h], e.g. OE hw-, e.g. hpeth WHAT, hpen WHEN;
8. ‘h’ following ‘p’ for historical hw-, e.g. phite WHITE, phile WHILE;
9. ‘h’ following ‘p’ for historical [w], e.g. unphrihen UNWOUND;
10. ‘h’ following ‘c’ in the digraph for historical [tf] and [ʃ], e.g. chirche CHURCH, flech FLESH, -chipe -SHIP;
11. ‘h’ following ‘c’ for historical medial and final [x], e.g. ibrocht BROUGHT, ah OUGHT;
12. ‘h’ following ‘g’ for historical medial [ʏ], e.g. daghes DAYS;
13. ‘h’ following ‘c’ for historical medial [ç] and [çt], e.g. michte MIGHT, richpisnesse RIGHTEOUSNESS (note also ‘chi’ in nichit NIGHT);
14. ‘h’ following ‘p’ for historical [f] in, phariseus gen. PHARISEE, prophet PROPHET (ultimately Hellenisms in Latin);
15. ‘h’ following ‘s(c)’ in the sequence for historical [ʃ], e.g. s(c)hulen SHALL pl., disch DISH;
16. ‘h’ in combination with ‘c’ and ‘s’ for historical final [ʃ], e.g. flechs, flecsh-FLESH, englichs ENGLISH;
17. ‘h’ following ‘s’ and in combination with ‘c’ for historical [sw], e.g. alsha ALSO, schuc SUCH;
18. ‘h’ following ‘t’ in the digraph for historical final [θ], e.g. deth DEATH, bi-holdeth BEHOLDETH;
19. ‘th’ for historical [t], e.g. geth YET (this word only, one example);
20. ‘h’ alone or ‘gh’ for historical final [θ], e.g. strecheth STRETCHETH (not in the LAEME tagged text: cited from Scahill 2008: note 17), druhieth DRY (3rd plural present indicative);
21. ‘ch’ for historical [k], e.g. flocch FLOCK, ach (< OE ac) BUT;
22. ‘nch’ for historical [ŋ], e.g. kanche-ship (of obscure origin but elsewhere normally cang-) FOLLY (note also unstrechpe UNSTRENGTH with missing ‘n’ or abbreviation);
23. ‘sth’ for ‘str’ in sthengðe STRENGTH (not in the LAEME tagged text: cited from Scahill 2008: note 17);
24. ‘soh’ for historical [ʃ] in i-sohouen SHOVED (this example only);
25. excrescent ‘h’ in hures OURS, huuel EVIL.

To say that the G scribe had 25 ‘different functions’ for ‘h’ would, however, be unparsimonious. The prodigality is not chaotic, but is systematically reducible. Examples (1)-(5) and (25) appear to be segmental. Of these, (1)-(3) continue traditional functions. Examples (4) (5) and (25) illustrate lenition processes; (4) and (5) may represent the result of lenition to [ʃ] and (25) probably represents Ø (see
§3.1 above). Examples (6)-(24) are most probably diacritic, implying fricativeness (for full commentary, including discussion of a similar set of ‘h’ spellings in a different text, see Laing and Lass 2009: § 5.2). So our 25 categories may be reduced to three: historical retention, lenition, fricative representation.

5. ‘h’ in Cambridge, Trinity College B.14.39 (323), Hand A
The weakness of [h] along with the common development of diacritic uses for ‘h’ makes for the very widespread appearance of the littera, whether alone or in combination(s). There are some curious ‘h’ spellings in the work of Hand A of Cambridge, Trinity College B.14.39 (323). Trinity Scribe A has some of the kinds of ‘h’ spellings detailed above for Scribe G of Ancrene Riwle. But there are also some we have not seen anywhere else, where ‘h’ appears intervocalically:

bahit and firbahit (beside bait, beid) for (for)BADE (3rd sg past < OE (for-)bād)
chehec for cheek (< OE cēace)
clohit (beside cloþis ) for CLOTHES (< OE clāh)
de’h’it (beside ded(e)) for DEATH (< OE dēaþ – with ‘h’ interlined by the scribe as a correction)
fehid and fehit (beside feid, feit and wed)¹² for FEET (< OE fēt)
gohid (beside god(e) and goid) for GOOD (< OE gōd)
maist1he for MASTERY with ‘i’ superscript standing for ‘ri’ (< AF maistrie)
nehic (beside ney) for NIGH (< OE nēah)
rothien for ROT (inf.) (< OE rottan)
wehit for PLEDGE (< OE wedd)

What does ‘h’ mean in these spellings? We could attempt separate explanations for each etymological form type above, but this would be unparsimonious. One explanation can suffice for all the spellings cited. Scribe A drops initial ‘h’ in non-lexical words, such as pronouns and auxiliary HAVE. He shows much more initial ‘h’ insertion – in 27 different lexemes not counting pronouns. We take ‘h’-dropping and insertion to imply that he had at least variable loss of initial [h], and that therefore ‘h’ can function as a null character as well as in other ways.

There are two unusual non-intervocalic rhyming ‘h’-spellings: behit : ireht for bed (< OE bedd) and read (past participle < OE gerǣd(e)d). The ‘t’ in these spellings suggests devoicing of the stem final consonant. Evidence is widespread that this had happened in Scribe A’s usage. Note e.g. bait BADE and hypercorrect feid FEET from the list above, as well as e.g. bont BOUND, breit BREAD, blint BLIND elsewhere

in his text, and hypercorrect spellings like *had* *that*, *hid* *it* and *vad* *what*. *beht : ireht* are different from these other examples in that they had stem-final historical geminates. The only similar spelling for the same etymological type is *wehit* (< OE *weddd*) listed above. We might postulate that the ‘h’ represents the first element of a devoiced geminate, as happened for instance in Icelandic. Pre-aspiration would be typologically odd for a dialect from this location, and would require more evidence. Since we do not know of any that would support this interpretation, and *wehit* with interposed ‘i’ rather complicates the idea, we postulate that the ‘h’ is null.

Many of the intervocalic ‘h’ in the list above are found in words for which Scribe A also uses digraph spellings involving ‘i’, but without ‘h’. There are many other such spellings in his text: e.g. *breit* *bread*, *feind* *fiend*, *for-soit* *forsooth*, *Goil* *gold*. The majority of these digraph vowel spellings in accented position represent long vowels, either ancient as in *for-soit* or more recent as in *Goil*. A few, e.g. *goid* *God* represent short vowels and some e.g. *afteir* *after*, *Eueir* *ever*, *wateir* *water*, occur in unstressed syllables. None of the digraph spellings represent sequences longer than one syllable. The intercalation of ‘h’ between the vowels cannot represent syllable junctions in historically monosyllabic words. Therefore, in these cases too, we take it that ‘h’ is a null character.

Can it be that this scribe simply enjoys variant untraditional but explicable spellings? Elsewhere (Laing and Lass 2003:261-262) we have discussed the fact that he employs four different spellings for OE -iht in rhymes in the same stanza of a poem – *brit : mist : vichit : nicst*, for BRIGHT, MIGHT, WIGHT and NIGHT. It is very difficult to think of any purpose for such apparently deliberate writing of non-identical sequences in rhymes than the celebration of potential variation.

University of Cape Town
University of Edinburgh

ROGER LASS
MARGARET LAING

REFERENCES


