

## TEMPLATE FOR THE SHAPE OF AN INDIVIDUAL CoNE ENTRY

The shape of an individual CoNE entry takes the form of entries to different sections in a structured template containing 13 fields:

1. TAG
2. DICTIONARY BOX
3. HYPERETYMOLOGY
4. CLASSIFICATION
5. OLD ENGLISH ETYMOLOGY
6. MIDDLE ENGLISH ETYMOLOGY – INTRODUCTORY NOTES
7. UNIQUE FORMS OF THE BASE ATTESTED IN LAEME
8. MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY
9. MIDDLE ENGLISH ETYMOLOGY – PROBABLE OLD ENGLISH INPUT PARADIGM TO MORPHOLOGY
10. MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY
11. DERIVATIONS
12. DERIVATION – COMPOUNDS
13. DERIVATION – GRAMMATICALISATIONS

The content of and conventions used in each of these fields is described below.

### 1. TAG

#### **General shape**

$\$[\text{lexel}]/\text{grammel} [< \$\text{lexel}/\text{grammel} \text{ plus link}]$

Further relevant links

Lexel = LAEME lexical element.

Grammel = LAEME grammatical element.

Material inside [...] is not always present.

Each etymology references under the TAG heading all the tags relevant to it. In the case of simplex items, no further information is provided under TAG. Any derivation (including compounds and grammaticalisations) is followed immediately by a link to the entry for the word from which it derives. Complex items (whether compounds or other prefixed or suffixed items) are provided with links to the etymologies of their parts. We follow standard dictionary practice and derive compounds from their leftmost lexical element. Links to each of the other elements of a complex item are also provided under TAG. Where the etymology of a complex

item is entirely covered by those of its constituents, the complex item is supplied with a DICTIONARY BOX entry but no further etymological narrative of its own.

## Further notes

### *Tags*

The tag identifies the item in the LAEME CTT for which CoNE provides a narrative etymology. Notationally, the most extended tag type consists of a lexical element ('lexel') and a grammatical element ('grammel'), which identifies part-of-speech and functional category. The lexel is introduced by \$ and the grammel by /. Not every item has a lexel: e.g. pronouns, determiners and inflectional affixes are not given lexical labels, because they can be construed as carrying only grammatical information. Some tags, therefore, may consist of a grammel only, but none of a lexel only. One kind of lexel provides exceptions to this rule in CoNE: there are occasional cases where the LAEME CTT has an attestation of a compound item but no attestation of its first element as a simplex: e.g. \$chalkstone/n for which there is one attestation, while LAEME lacks any instances of the word CHALK alone. In this case the CoNE lexel is \$chalk~, with no / or following grammel, since in the context of LAEME CTT the /n extension properly belongs not to \$chalk but only to \$chalkstone. The lexel for AND is \$and/cj corresponding to LAEME \$&. Lexels for numerical quantifiers are arabic numerals with the grammel specifying whether cardinal /qc or ordinal /qo. Lexels for derivational suffixes begin with -, e.g. \$-less/xs-aj.

### *Lexel types*

Lexels are of five basic kinds: 1 Modern English; 2 Old English; 3 Scandinavian; 4 Middle English; 5 composite (see further under DICTIONARY BOX below).

Where two lexels (or two lexeme-defining grammels) are in a co-ordinating construction the LAEME CTT their grammels are extended by >= and <= respectively, e.g. \$neither/cj>= and \$nor/cj<=, \$so/av>= and \$as/cj<=. These tags and their forms in LAEME CTT are normally subsumed under the unextended tag, e.g. \$neither/cj.

### *Conventions for Old English*

The only Old English characters appearing in LAEME lexels are aesc, edh and thorn. These are indicated respectively by ae, D and Y. Long vowels are indicated by a following colon not by a macron on the vowel.

### *Specifiers*

Some lexels are followed by annotations in braces, which are specifiers usually indicating some semantic or functional distinction: e.g. \$about{p}/pr, \$about{re}/pr, \$about{t}/pr identify the prepositional lexel \$about when referring respectively to PLACE, CONCERNING and TIME. For a full list see the CoNE document entitled LAEME Lexel Specifiers.

### *Grammels*

For full details see the separate CoNE document entitled Grammel Commentary.

Basic grammels that are part-of-speech identifiers (always following a lexel): aj = adjective, av = adverb, cj = conjunction, n = noun, pr = preposition, v = verb.

For extensions to the v grammel that give further grammatical information see commentary under OLD ENGLISH ETYMOLOGY, **Further notes**, *Items with Ablaut or i-umlaut*.

Basic grammel-only tags:

## 1. Determiners, pronouns and articles

A = indefinite article, A

Dat = determiner, THAT

Des = determiner, THESE

Dis = determiner, THIS

Dos = determiner, THOSE

P = personal pronoun

RTA = relative pronoun animate (normally human)

RTI = relative pronoun inanimate or non-human (unless the non-human is treated as a person)

T = definite article THE

## 2. Inflexional affixes

Many grammels include further specified subsets.

/Gaj = genitive inflexion on adjectives

/Gn = genitive inflexion on nouns (includes also) \$/Gpn

/Odaj = in adjectives, invoked only with formal survival of a reflex of OE acc. masc. *-ne*

/plaj = plural inflexion on adjectives

/pln = plural inflexion on nouns

/v-imp = verb imperative sg ending

/v-imp22 = verb imperative pl ending (includes also \$/v-imp02)

/vSpp = verb strong past participle ending

/vSpt2 = verb strong past tense plural ending

/vi = verb infinitive unmarked ending, excluding weak class 2

/vi-m = verb infinitive marked ending, excluding weak class 2

/viK2 = verb infinitive weak class 2 unmarked ending

/viK2-m = verb infinitive weak class 2 marked ending

/vn = verbal noun ending, excluding weak class 2

/vnK2 = verbal noun ending, weak class 2

/vpp = weak verb past participle ending, excluding weak class 2

/vppK2 = weak verb past participle ending, weak class 2

/vps11 in verbs excluding weak class 2, invoked only if the ending is other than E

/vps11K2 = present indicative weak class 2 1sg ending

/vps12 = present indicative, excluding weak class 2, 2sg ending

/vps12K2 = present indicative, weak class 2, 2sg ending

/vps13 = present indicative, excluding weak class 2, 3sg ending

/vps13K2 = present indicative, weak class 2, 3sg ending

/vps2 = present indicative, excluding weak class 2, pl ending

/vps2K2 = present indicative, weak class 2, pl ending

/vpsp = present participle ending, excluding weak class 2

/vpspK2 = present participle ending, weak class 2

/vpt1 = past tense 1sg and 3sg ending, excluding weak class 2

/vpt1K2 = past tense 1sg and 3sg ending, weak class 2

/vpt12 = past tense 2sg ending, excluding weak class 2

/vpt12K2 = past tense 2sg ending, weak class 2

/vpt2 = past tense plural ending, excluding weak class 2

/vpt2K2 = past tense plural ending, weak class 2

13 August 2013

### 3. Negation

The tag for the negative particle *ne* and its variants is `$/neg` which is invariably further extended to indicate more specific syntactic information in LAEME but not relevant to CoNE.

For extensions on the grammels that give further morpho-syntactic information dealt with in MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY see the commentary under that heading below and for full details see the separate Commentary on the Grammels.

#### *Ordering of Tags*

The lexels in CoNE are ordered alphanumerically as in the LAEME text dictionaries, though the leading `$` is stripped out in the CoNE index. The alphanumeric ordering is a function of the computer methodology employed. This computational ordering means that symbols are usually listed first so that derivational and degree suffixes are the first set of elements listed under `-` (e.g. `-dom` etc). Then come the grammel-only tags that define lexemes under `/`, e.g. `/P13NM HE`, `/RTA WHO, THAT`. (Note that `~` however comes after `Z`.) `-` and `/` are followed by the numerals (including numeral compounds), and then comes the alphabetical list of ‘normal’ lexels. Here too, however, the ordering is not entirely as in a standard dictionary because thorn precedes `A`. In the LAEME text dictionaries, initial capitals are automatically listed first, which means that capital `Y` used for Old English lexel-types beginning with thorn (e.g. `$Ya:` etc) come before `$a-` etc. This precedence of capital over lower case letters applies also within a lexel — capital `D` for `edh`, and `Y` for thorn — which can skew the expected ordering within the alphabetic listings also within CoNE. Additionally, within each letter’s listing the colon length symbol takes precedence over any letter, so e.g. `a:gan/v` comes before `aback/av`; in fact all long vowels anywhere within a lexel take precedence over the equivalent short vowel in the same alphabetical sequence. Unless the *ge-* form is the only one attested, Old English verbs with *ge-* prefix are normally combined with the equivalent verb without prefix under the prefixless form. Other prefixed verbs have their own lexels.

## 2. DICTIONARY BOX

### **General shape**

`$/[lexel]/grammel, DEFINITION, links.`

When the tag includes a lexel, it may be any of the following types:

1 Modern English equivalent; 2 Old English etymon; 3 Old Norse etymon; 4 (where none of 1-3 is apposite) the equivalent MED headword; 5 composite: Old English and Modern English.

There follows a basic definition in small capitals with or without other explanation in Roman. The definitions do not seek to be exhaustive, but they do aim to cover the main usages found in the LAEME CTT.

Beneath are links to the relevant entries in the online versions of OED and/or MED and (where available) DOE, which at the time of writing only goes up to `G`.

### Further notes

The etymology of a complex item (whether a compound or other prefixed or suffixed item) may be entirely covered by those of its constituents. Such an item is supplied with a DICTIONARY BOX entry but no further etymological narrative of its own. In these cases, links to the etymologies of the constituent simplexes are given under the TAG heading.

### 3. HYPERETYMOLOGY

This section is intended for a later phase of CoNE.

#### General shape

**\*superordinate etymon** with link.

Where relevant, a link is provided to a separate HYPERETYMOLOGY entry. This gives the etymological pathways for each member of a group of items attested in the LAEME CTT that are related underlyingly only by means of descent from a parent form earlier than our usual input form. Each member of the group is identified by its LAEME tag (see above under TAG) and each has a link under this heading to indicate the etymological relationships between itself and any derivatives it might have and the other items and any derivatives they might have, some of which may also have derivatives. Links from nodes in the hyperetymology lead to the narrative etymology (CoNE entry) of individual items. A hyperetymology will typically involve divergent pathways that originate earlier than primitive Old English and/or Old English and Old Norse reflexes of the same base. (For the definition of what is covered by the term ‘base’ in CoNE, see below under UNIQUE FORMS OF THE BASE ATTESTED IN LAEME.) The narrative will therefore have a ‘back story’ that predates our usual entry point. For instance, the hyperetymology of **\*full-** will delineate the historical narrative that accounts for the derivational relationship between the development of the items: \$full/aj, \$fultum/n, \$fulluht/n, \$fill/v, fylstan/v, \$follow/v, \$follow/vK2, including all their own (sometimes numerous) derivations.

### 4. CLASSIFICATION

Nouns: strong, weak, gender and (where available) stem classes.

Pronouns and determiners: gender, number, case/function, — as indicated in the grammel for each.

Verbs: strong, weak and class.

Adjectives: utilise endings from nominal stem declensions as specified. They may decline strong or weak.

### 5. OLD ENGLISH ETYMOLOGY

#### General shape

13 August 2013

|| **reconstructed ‘input’ form ((change))** > **\*resulting form** > [phonetic representation of the OE citation form] > *OE citation form*

|| introduces the etymology. The input form is normally a reconstruction of the item in the form we suppose it to have had at the time of the Anglo-Saxon Settlement. Reconstructed forms are indicated by asterisk and bold font. In some cases the entry point has to be from an earlier stage in order to account for (all of) the item’s subsequent history. In one particular case, we reconstruct a segment, which we know had disappeared as early as West Germanic. This is the stem-formative **\*-a** in the masculine and neuter *a*-stem nouns, which is anachronistically included for pedagogical purposes. Although this stem class marker had long since disappeared, we include it in our reconstructed Proto-Old English form to indicate the reason for the assignment of this particular class. We then remove it as early as possible in the temporal sequence via the change Weak Final Vowel Deletion ((WFVD)).

Old English etymologies may have Pathways and/or Forks. Pathways are labelled (A), (B), etc. and Forks (which by definition branch off from Pathways) are labelled ①, ②, etc. Labelled Pathways are invoked when there are two or more entry forms of different shapes; the default etymology is a single Pathway, with or without Forks. A Fork is a later differentiation in a Pathway.

If there is more than one reconstructed entry form necessary to account for the item’s subsequent history, || introduces a ‘back story’ form — a further abstraction of the reconstruction that can account for both entry forms, after which the etymology then proceeds down separate Pathways labelled (A), (B), etc. E.g. in the etymology of *-er/xs* (the suffix indicating comparative) there are two entry pathways: (A) **\*-ir-a** and (B) **\*-o:r-a** for which the superordinate ‘back story’ form is ||**\*-Vr-a** where V stands for either vowel.

((change)) refers to a set of bracketed initials identifying a change in the Corpus of Changes to which it is linked and where a description of the change is found. > represents ‘becomes as a result of the preceding change’. There may be multiple ((change)) > and **\*resulting form** stages. After any > there may also be a Fork, marked ①, ②, etc. to account for differing outputs and eventual *OE citation forms*. The ideal or standard relation of changes is that each pair should be representable in the form ((change)) > ((change)) with the > indicating a time axis. (For discussion of relative chronology of changes see Introduction, §10.5.) However, in certain circumstances changes may be bundled. Where the order of any pair or n-tuple of changes is undecidable, they are linked by &, e.g. Analogical Levelling & Weak Vowel Neutralisation – ((AL)) & ((WVN)) – for any given paradigmatic ending. Where any pair of changes are linked by a comma (,) the comma signifies ‘then’. This can occur, for instance, in situations where one change serves as a well-formedness condition on the output of the other, e.g. Degemination, Medial Fricative Voicing — ((DG)), ((MFV)) — where the single fricative produced by the first change is immediately voiced by the second change. Where the ordering of changes to stems or inflectional endings is not decidable, we generally allow changes to operate first on lexical material and then on grammatical material. (For the significance of linking comma vs linking & in the Middle English part of an etymology see below under MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY.)

Reconstructed forms (asterisked in bold) or forms in square brackets are phonetic

representations at a very broad ‘typological’ level of transcription. Short vowels are unmarked, long vowels are indicated by a colon. Long diphthongs are unmarked and the first elements of short diphthongs are marked with breves, following the conventions in Hogg (1992a) and Lass (1994). The form in square brackets immediately preceding the Old English citation form is always the phonetic equivalent of the citation form.

Old English citation form(s) at the end point(s) of an etymology serve as the entry points to the BASE PHONOLOGY section of the MIDDLE ENGLISH ETYMOLOGY that follows. Old English forms are given in italics, following traditional practice: short vowels and diphthongs are unmarked, while long vowels and the first element of long diphthongs are marked with macrons. The Old English forms given are the most frequently attested variants and are not intended to be exhaustive. Occasionally a minor variant, as listed in DOE or Bosworth-Toller, may be included if it is plausibly a source for a particular spelling in the LAEME CTT that is otherwise difficult to account for. We do not normally list variants from DOE or from other sources if they are attested only in post-Conquest texts. It must be remembered however, that given the diachronically and diatopically patchy survival of both Old English and early Middle English texts, it would be dangerous to assume that a single unusual Old English attestation was necessarily the direct source of a particular early Middle English variant. We follow Old English handbook and dictionary practice in retaining the Old English characters ‘æ’, ‘ð’ and ‘þ’ but changing runic ‘p’ to ‘w’. We also print OE ‘ġ’ as ‘g’.

For Old Norse loans, we also provide citation form(s) in italics, which then serve as entry points to the BASE PHONOLOGY section of the MIDDLE ENGLISH ETYMOLOGY of any LAEME CTT items or forms that cannot be accounted for by origins in native Old English. Even when such forms are unattested in the Old English record, we nevertheless treat them as part of the OLD ENGLISH ETYMOLOGY, because their presence in Middle English suggests earlier borrowing.

## Further notes

### *Implications of citation practice for ‘w’ and ‘g’*

The changing of OE ‘p’ to ‘w’ and ‘ġ’ as ‘g’ in the citation forms makes for continuity with usage in other sources, but leads to an anomalous situation when we proceed to the MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY section for items that contain these *litterae*. In Middle English, there emerged new *litterae* that gradually took on functions previously played by Old English letters or lettershapes. The *littera* ‘w’, which developed from ligatured double ‘v’ / ‘u’ became an alternative representation for [w] alongside ‘p’, see ((EOW)). Caroline minuscule script (earlier reserved for writing Latin texts) provided a new Caroline *g figura* which took over the functions [g] and [dʒ] while insular ġ retained the [j] and [ɣ] functions, see ((EOCG)). Early Middle English texts to varying degrees preserve ‘p’ and in LAEME CTT it is differentiated from the later *littera* ‘w’: ‘p’ is realised in LAEME as lower case w and ‘w’ as upper case W. Similar differentiation is made in LAEME CTT between the different *figurae* for ‘g’: Caroline ‘g’ (*g*) is realised as upper case G, surviving insular ‘g’ (*ġ*) as lower case g. The description of the Special Code (for Special Codes see below under MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY) ‘No Change from Old English’ ([NCOE]) takes this anomaly into consideration and allows for both w ‘p’)

and W ('w') spellings and both g ('ġ') and G ('g') spellings to be marked ([NCOE]) in the MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY section of the etymology of any relevant item. In such cases, the MIDDLE ENGLISH ETYMOLOGY – INTRODUCTORY NOTES SECTION makes it clear that the change ((EOW)) or ((EOCG)) is to be applied to all examples of W or G in the LAEME CTT attestations. This obviates what would otherwise be the clumsy necessity of assigning the change individually to every W or G in the LAEME CTT forms cited. (For explanation of the LAEME transcription policy for other Old English letters see below under UNIQUE FORMS OF THE BASE ATTESTED IN LAEME.)

#### *Items with Ablaut or i-umlaut*

The Old English etymologies of some items have multiple starting points, which are equivalent to Pathways of the kind described above, but which for reasons of clarity are labelled differently. Items of this kind are strong verbs and nouns that show i-umlaut in some paradigm members. Ablaut series in the various classes of strong verbs result in principal parts with different vowels (though not all classes have different root vowels in each part). For strong verbs we give separate etymologies for each of the principal parts, each of which results in Old English end points. These in turn provide the inputs to the narratives in the BASE PHONOLOGY section of the MIDDLE ENGLISH ETYMOLOGY that follows. Nouns such as \$foot/n, \$man/n are treated similarly, with unumlauted and umlauted forms being treated as two different 'principal parts'. The grammels introducing the strong verb principal parts are: \$/vi (verb infinitive); \$/vSpt1 (verb strong past sg); \$vSpt2 (verb strong past part); \$vSpp (verb strong past participle).

#### *Demotivated compounds*

Etymologies of compounds that have already gone through reduction processes, e.g. *gospel* < *good spell*, begin with the compound elements already in their complete OE form. The etymologies of the simplex elements are linked.

#### *Treatment of derivations*

Derived items are given a separate Old English etymology only in cases where the narrative is different from that of their source: e.g. the Old English etymology of \$good/aj provides the etymology also for derived \$good/av and \$good/n. Such derivations are given a CoNE entry under the TAG and DICTIONARY BOX headings and are cross-referenced (as derivatives) for the rest of their narrative to the etymology of \$good/aj: e.g. the TAG heading entries are \$good/av < \$good/aj and \$good/n < \$good/n respectively.

Compound items, including those made up of a base plus derivational suffix, are not normally given their own separate Old English etymologies; the etymology of each element is dealt with under the respective simplex etymology (see further under DERIVATIONS and DERIVATION – COMPOUNDS below). An item of this kind is given a CoNE entry under the TAG and DICTIONARY BOX headings and cross-referenced as being derived from the simplex of the first element and further cross-referenced to the etymologies of both elements that make it up: e.g.

\$dre:amYyrl/n < \$dre:am/n

For the first element see the etymology of \$dre:am/n.

For the second element see the etymology of \$Yyrl/n.

If a derived item has no Old English attestations and is generally accepted as being a Middle English formation, it is not given a separate narrative etymology. The CoNE entry for such an item provides information only under the TAG and DICTIONARY BOX headings and its etymology is dealt with in the etymology of the ‘parent’ form under DERIVATIONS, where further links are given if appropriate.

A derived item which has an independent Old English existence and for which the base has developed differently from that of its ‘parent’ is given its own full etymological narrative in CoNE. Some other derived items, e.g. all denominal verbs, are also given their own narrative etymology even if the etymology of the base is identical with that of the base of the parent item. This is for the very practical reason that a narrative is needed also for the verbal morphology of the item. The complex Old English verbal paradigm that provides the input forms for the MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY section is better listed in a separate entry simply for the sake of clarity. So, for instance \$bridge/v is listed as a derivation under DERIVATIONS in the etymology for \$bridge/n, but from there is linked directly to its own full separate etymology.

## 6. MIDDLE ENGLISH ETYMOLOGY – INTRODUCTORY NOTES

This is not always be invoked but can include any general introductory remarks or specific explanations about the treatment of the Middle English part of the narrative etymology in question. Where relevant it includes invocation of one or more of the changes ((EOW)), ((EOCG)) or ((EOQ)) where these apply to all LAEME citations of forms showing W, G or Q respectively in the following BASE PHONOLOGY SECTION, and where it would therefore be clumsy, and/or confusing to the narrative of other changes, to invoke them separately for each citation.

## 7. UNIQUE FORMS OF THE BASE ATTESTED IN LAEME

The term ‘base’ is taken to mean the substance of a linguistic item minus any inflectional morphology. Thus the base of a free morpheme is the same as its root. By extension we use it also to cover the substance of items that are bound morphemes, e.g. derivational or inflectional affixes. Here are given all citations of the base of the item in the LAEME CTT as listed in the LAEME Tag Dictionary. The forms are cited in LAEME internal format: plain text manuscript letters are realised as upper case, lower case letters are employed to represent (a) the traditional expansions of manuscript abbreviation sigla and (b) the Old English ‘special’ letters viz. ae = aesc, d = edh, g = insular ‘g’, w = wynn, y = thorn, z = yogh. Manuscript capitals (majuscule) and *litterae notabiliores* are not differentiated from manuscript minuscule letters in CoNE other than in the 1sg pers pron. *ī* where capitalisation has orthographic developmental significance. In this one case capital *ī* is realised as in the LAEME CTT with a preceding asterisk: \**ī*. Acute accents occasionally found on vowels are indicated with lower-case x after the accented letter, e.g. manuscript *á* is realized as Ax. Superscript letters in the manuscript are preceded by ^. Inserted letters in the manuscript are transcribed between > and >. Deleted letters in the manuscript, where in special circumstances they are retained, are transcribed between < and <.

In this section we identify all unique forms of the base, i.e. the unique spellings in LAEME CTT. In the case of items, which are free morphemes, this means once any and all inflectional and derivational affixes are removed. These spellings are input to the MIDDLE ENGLISH

ETYMOLOGY – BASE PHONOLOGY section. A leading or trailing +, - or ± indicates that the form in question is attested only in combination with some other morpheme. (In LAEME format, - indicates that in the manuscript the elements are separated by a space, while + indicates that there is no space between the elements in the manuscript text. ± is a CoNE convention used to indicate that the elements it divides are not treated as separate morphemes in LAEME.) While a spelling in final E could indicate an inflected form (more likely in early than in late texts), it need not do so. For example, in LOUE IS SWET LOVE IS SWEET (LAEME CTT text # 2002 digby86mapt.tag), the final E of LOUE no doubt answers historically to *-u*, the OE nom sg inflection of SF *ō*-stem nouns like *lufu* LOVE. But whether it functions as a grammatical morpheme signifying nom sg or whether it has come to be regarded as part of the base is unknowable, and the answer may differ from scribe to scribe. Because of the deterioration of the case system in ME, pairs such as BAy and BAyE could instead indicate mono- and disyllabic forms of the same word, with final E no more functional than the ‘e’ at the end of ‘programme’ or the ‘a’ at the end of ‘china’. Because of our uncertainty about the status of weak final vowels, we list as unique any form that is distinguished from another only by the presence of such a vowel, i.e. we list BAyE as well as BAy as a unique form of \$bath/n. Any commentary on the quality and potential significance of weak final vowels (typically, but not always, in E or e) is, however, confined to the MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY section.

## 8. MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY

### General shape

#### *OE citation form*

[[((change(s))) >] [[code(s)] >] LAEME BASE FORM(S)

#### Notes

Repeat of change(s) and or special code(s) resulting in different LAEME base forms, and accompanying notes, until all developments are accounted for. Special codes are employed to identify certain linguistic phenomena that are not in themselves changes but are nevertheless part of the story of the form shapes recorded in LAEME. Each such phenomenon is identified by a short abbreviation. To distinguish these phenomena from those described in the Corpus of Changes, we give special codes in brackets within parentheses, e.g. ([NCOE]) for No Change From Old English, ([NPTR]) for Northern Present Tense Rule.

For each Old English input form in italics we deal with the descendant LAEME base forms (in LAEME internal format), as listed under UNIQUE FORMS OF THE BASE ATTESTED IN LAEME. All post-Old English developments in the shape of the base are accounted for by reference to one or more changes, denoted by double-bracketed initials and described in the Corpus of Changes (CC) or to one or more special codes, denoted by initials brackets within parentheses and described in Special Codes. > represents ‘becomes as a result of the preceding change’ or ‘is/remains with reference to the preceding code’. Abbreviations and corrections in manuscript spellings are also treated under this heading. Expansions of abbreviations are ‘deemed to represent archetypal’ categories. The reason for this form of words is that abbreviation symbols do not have phonological content and hence cannot

properly be said to map into segments. However, spellings that include abbreviation symbols still require to be accounted for in the phonological narrative. Forms with final weak vowels are listed alongside those without, but commentary on them is in the MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY section, even where the final vowel may be considered to be part of the base, as with the final vowel of e.g. NAME (< OE nom. sg. *nama*).

When changes or codes are bundled, linking comma implies that they occurred in temporal sequence, linking & implies that the relative timing is undecidable.

See above under Old English Etymology, **Further notes**, *Implications of citation practice for 'w' and 'g'* for an explanation of the mismatch between Old English citation forms in *w* and *g* and LAEME CTT citation forms in *W* and *G*; cf. also ((EOW)) and ((EOCG)).

## **Futher notes**

### *Chronology*

There is a gap in the written record between written Old English and early Middle English. The gap is very partially filled by 'transitional English' in the form of late C11 to late C12 copies of Old English works, to a greater or lesser extent 'updated'. LAEME CTT includes such texts only when (a) they are from manuscripts dated later than ca. 1150 and (b) their language shows significant change towards early Middle English. Even if the narrative provided by transitional English were to be included in CoNE, evidence of direct descent from Old English paradigms to eME output forms would be very patchy. The developments in the base dealt with under MIDDLE ENGLISH ETYMOLOGY, and referred to above as 'post-Old English', will (by definition) have occurred before the LAEME attestations. Some of them, however, may have already happened in some forms of Old English, which either are unattested or are attested in less common citations that we have not listed in the OLD ENGLISH ETYMOLOGY section that provides our input form(s). Here we take the terms 'post-Old English developments' to include some late Old English and 'transitional' developments and MIDDLE ENGLISH ETYMOLOGY loosely to refer also to those periods where relevant.

## 9. MIDDLE ENGLISH ETYMOLOGY – PROBABLE OLD ENGLISH INPUT PARADIGM TO MORPHOLOGY

The Old English paradigms given in any etymology follow the standard ones given in the handbooks. These are for the most part West Saxon, which is problematic, as most of our surviving Middle English texts are not of West-Saxon provenance. We follow this practice because these are the most familiar paradigms, and are taken at large to define 'Old English morphology'. They are conventional rather than strictly genetic paradigms (which given the amount of variation in Old English would not strictly speaking be obtainable anyway); but they serve the purpose of outlining the forms most commonly referred to in the literature. So we list the endings of a normalised historical paradigm. In cases where there is disagreement in the literature about morphological paradigms, we adopt a compromise position and generally follow the versions given in Campbell (1959).

It may seem curious that  $-\emptyset$  is given as a possible ending in some cases.  $-\emptyset$  is not to be taken as a morph, the way e.g. *-um* would be, but as a descriptor and a sign of the results of history.

To say that a form has a zero ending is not to say that it has a zero morph, which is a technical device used in certain kinds of synchronic linguistic analysis as something that can be operated on, replaced, etc. This zero is purely an indication that the word-form in question is endingless, for the most part through some process(es) of deletion that took place during its early etymological development. Thus the  $-\emptyset$  in the nom sg of *a*-stem nouns means only that before attested Old English the nom sg ending was deleted, leaving a bare root to represent that case/number form.

### **General shape**

Applicable only for nouns, adjectives and verbs.

For a noun: a paradigm showing the morphological endings relevant to the Old English classification, including sg and pl for all grammatical cases. For nouns of the mutation declensions the base is also included.

For an adjective: paradigms showing the morphological endings for all three genders, sg and pl, strong and weak.

For a verb: the full paradigm for the specified verb, including the base.

The paradigms or paradigmatic endings are supplied to provide points of reference from which the later morphological developments can be delineated. Verb paradigms to a great extent maintain conjugational differentiations between Old and Middle English, so the etymological narrative is for the most part transparent. For nouns and adjectives, however, distinctions were largely eroded by Middle English times. Nevertheless, we supply both weak and strong paradigms for adjectives, so that clearly sourced reflexes can be interpreted by reference to them. It is, however, most often the case that the form in LAEME CTT is uninformative as to whether it belongs to an original weak or strong paradigm. See further under the heading: MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY below. Although comparative and superlative adjectives are in one sense part of the adjectival paradigm, they are not part of the inflectional paradigm because they themselves inflect. The ‘-er’ and ‘-est’ suffixes are given their own etymologies in CoNE (like other derivational suffixes, such as ‘-isc’ and ‘-ly’). Their inflectional morphology is therefore dealt with under these separate suffixal etymologies, viz \$-er/xs-cpv and \$-est/xs-sup.

## 10. MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY

### **General shape**

\$([lexel])/grammel. Links to etymologies of separable morphological affixes where relevant.

/grammel(s)

[((change(s))) >] BASEFORM+ENDING(S)

## Notes

Repeat of all three latter stages, i.e. /grammel(s), change(s) resulting in different LAEME morphological endings, and accompanying notes, until all grammels and the morphological developments of the forms belonging to them are accounted for. A change may not always be invoked. In such cases the accompanying notes offer explanation.

For full details on grammels, see the separate Commentary on the Grammels.

Grammels for nouns, adjectives and pronouns have the following extensions indicating case/grammatical function (no extension implies subject/nominative): Od = direct object, G = genitive, Oi = indirect object, pr = prepositional object. To indicate the position of a form relative to a governing preposition, the pr element is preceded either by < (indicating the pronoun follows the preposition) or > (indicating the pronoun precedes the preposition). When a noun functions as an adjective its grammel is \$/naj and when an adjective functions as a noun its grammel is \$/ajn. These grammels may be further extended as above.

Grammels for verbs have extensions indicating tense, mood, number, and person. We assume a two-tense system: vps = verb present, vpt = verb past/preterite. Verb grammels are built up in the same way as they are for other categories, as concatenations of features. The order chosen is number followed by person: e.g. \$/vps13 = verb present singular third person; \$/vpt21 = verb past plural first person. \$/vi = infinitive, \$/vi-m = marked infinitive (preceded by a form of TO), \$/vn = verbal noun (gerund), \$/vpsp = present participle, \$/vpp = weak past participle. S indicates 'strong' in past tense and past participle: \$/vSpt13 = strong past 3sg, \$/vSpt22 = strong past 2pl, \$/vSpp = strong past participle. \$v-imp = imperative singular, \$v-imp-22 = imperative plural. sj indicates subjunctive: e.g. \$/vsjps12 = present subjunctive 2sg, \$/vsjpt13 = past subjunctive 3sg. The sj extension is only invoked where the LAEME spelling is formally distinct from that of the indicative form of the same number and person. The grammels of verbal nouns are given noun extensions as appropriate. When participles are used as attributive adjectives their grammels are given the extension -aj and further adjective extensions as appropriate, e.g. \$/vpp-ajplOd.

All morphological endings (including zero) in the LAEME CTT are accounted for under their respective grammels, regardless of formal identity. Some grammels, however, may be amalgamated as equivalent; e.g. forms listed under grammels with the extension -k (indicating second element of a compound), or -t (in nouns indicating a title, e.g. \$earl/n-t) or {rh} indicating rhyme context, are subsumed under the relevant simplex grammel. The material is normally presented in alphanumeric order of the grammels, apart from subsumed equivalent categories. In the Middle English morphology section of the etymologies of strong verbs, however, the grammels are grouped under the headings of the Pathways identified in the base phonology section for the verb's principal parts. Within each of these groupings, the grammels are presented in alphanumeric order. Morphological endings that include, or take the form of, abbreviations are also treated under this heading, as are endings with scribal corrections.

LAEME CTT grammels do not exactly replicate Old English case labels. Od may in most instances be taken to label reflexes of old acc, and Oi reflexes of old dat. <pr and >pr labels, however, may attach to reflexes of a number of different cases, and usually it is not knowable from the reduced form of the reflexes, which case it might be. The most common post-prepositional cases in Old English were dat and acc, the latter in many noun classes zero-

marked in the singular, the former in *-e* or *-an*. In the plural the endings for acc would have been *-an*, *-as*, *-a*, *-e*, *-u* or *-∅* and for the dat *-um*. By Middle English times these distinctions were largely eroded, and (except perhaps on metrical grounds in the case of some verse texts), it is not determinable what any noun ending of a simple vocalic shape might have meant in terms of the ‘old’ cases. By about the 11th century we can assume one non-genitive ‘oblique’ case for post-prepositional nouns, and loss of the semantic distinction of dat/acc. Morphological commentary recognises this.

For an adjective, the form in LAEME CTT is usually uninformative as to whether it belongs to an original weak or strong paradigm. In those cases it is not feasible to check every instance of the word shape in LAEME CTT to determine the syntactic context, and it is left to the users to check the contexts for themselves via the LAEME website. In undetermined cases such as these, this phrase is commonly invoked: ‘whether weak or strong may only be determined by syntactic context in individual cases’.

Adjectives are dealt with under three headings: **Positive** (whose forms are dealt with under this heading); **Comparative** and **Superlative** are treated under DERIVATIONS.

The morphological endings of major category prefixed items are dealt with under their own entries under DERIVATIONS.

## 11. DERIVATIONS

We distinguish three categories of derivation. Compounds and grammaticalisations are, for procedural reasons dealt with under their own headings: DERIVATION – COMPOUNDS and DERIVATION – GRAMMATICALISATIONS below. Under this unspecified heading come all other kinds of derivatives: items formed from a base plus one or more derivational affix and all denominal, deadjectival and deverbal formations that are not categorisable as either compounds or grammaticalisations.

### **General shape**

Opening instruction, viz:

For phonological properties of base see MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY above. For morphological properties of derivation(s) see below or, where indicated, see the etymology of the suffix for any inflectional endings.

For etymologies of adjectives only:

### **Comparative**

### **Superlative**

For any etymology (including adjectives if there are any derived forms other than comparatives or superlatives):

⇒ \$lexel/grammel of derived item.

[Commentary as per opening instruction or as in MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY.]

⇒ indicates ‘derivation’.

Material inside [...] is not always present.

The sequence from ⇒ to the end of the commentary for that derived item is repeated for however many derivations from the parent form are attested in the LAEME CTT.

Where a derived item has been given its own etymological narrative in CoNE the entry under DERIVATIONS takes the form of a link to the item’s own CoNE entry.

## 12. DERIVATION – COMPOUNDS

### General shape

Opening instruction, viz:

For phonological properties of base see MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY above. For morphological properties of base (if any) see below.

⇒ \$lexel/grammel of compound item. For second element see \$lexel/grammel (linked)

⇒ indicates ‘derivation’. The sequence from ⇒ to the end of the commentary for that compound item is repeated for however many derived compound items are attested in the LAEME CTT.

For compounds, only the etymology of the matching simplex first lexical element is dealt with, i.e. complex items in which the first element is a prefix are dealt with under DERIVATIONS. The first lexical element of a compound is normally uninflected. The default comment on its morphology is ([CNM]) ‘Carries no morphology’. Links are given to and from non-initial element entries, which are dealt with under the etymology for their own simplex.

## 13. DERIVATION – GRAMMATICALISATIONS

Opening instruction, viz:

For phonological properties of base see MIDDLE ENGLISH ETYMOLOGY – BASE PHONOLOGY above.

⇒ \$lexel/grammel of grammaticalised item.

[Commentary as per opening instruction or as in MIDDLE ENGLISH ETYMOLOGY – MORPHOLOGY.]

⇒ as above indicates ‘derivation’ but in this section, specifically ‘grammaticalisation’.

Material inside [...] is not always present.

The sequence from ⇒ to the end of the commentary for that grammaticalised item is repeated for however many grammaticalisations derived from the ‘parent’ form are attested in the LAEME CTT.

Grammaticalisations are lexical items that have become delexicalised to form prepositions, prefixes or suffixes, etc. \$-dom/xs-n < \$doom/n or \$for-/xp < \$for/pr.