

# Medieval Text Types and the Corpus of Old English



P. S. Langeslag



## General Characterization

# Historical Characterization

## Prior to the Twelfth Century

- ▶ Church and state controlled the means of literary production

## Twelfth Century and Beyond

- ▶ Universities
- ▶ A secular bureaucratic class
- ▶ Eventually, the printing press

# Selected Written Languages in Northwestern Europe, and Gothic

## Prior to the Twelfth Century

- ▶ Latin
- ▶ (Greek)
- ▶ Gothic
- ▶ Old English
- ▶ Old High German
- ▶ Old Saxon
- ▶ Old Welsh
- ▶ Old Irish

# Selected Written Languages in Northwestern Europe, and Gothic

## Prior to the Twelfth Century

- ▶ Latin
- ▶ (Greek)
- ▶ Gothic
- ▶ Old English
- ▶ Old High German
- ▶ Old Saxon
- ▶ Old Welsh
- ▶ Old Irish

## Twelfth Century and Beyond

- ▶ Latin
- ▶ (Greek)
- ▶ Middle English
- ▶ Middle High German
- ▶ Middle Low German
- ▶ Old Frisian
- ▶ Old Norse
- ▶ Middle Welsh
- ▶ Early Modern Irish

# Historical Characterization: Text Types

## Prior to the Twelfth Century

- ▶ Scripture (mostly Latin, Gothic)
- ▶ Liturgical manuscripts
- ▶ Theological commentaries and treatises
- ▶ Royal law codes
- ▶ Canon law and monastic rules
- ▶ Homilies
- ▶ (Religious) historiography and annals
- ▶ Learning (grammar, computistics, arithmetic geometry, cosmology, medicine, classical poetry)
- ▶ Letters
- ▶ Records (mostly charters)
- ▶ Poetry (mostly religious, some lyric or heroic)

# Historical Characterization: Text Types

## Prior to the Twelfth Century

- ▶ Scripture (mostly Latin, Gothic)
- ▶ Liturgical manuscripts
- ▶ Theological commentaries and treatises
- ▶ Royal law codes
- ▶ Canon law and monastic rules
- ▶ Homilies
- ▶ (Religious) historiography and annals
- ▶ Learning (grammar, computistics, arithmetic geometry, cosmology, medicine, classical poetry)
- ▶ Letters
- ▶ Records (mostly charters)
- ▶ Poetry (mostly religious, some lyric or heroic)

## Twelfth Century and Beyond

Same, plus:

- ▶ Romance
- ▶ Drama
- ▶ More vernacular scripture
- ▶ More records: tax, chancery, exchequer, etc.

Much of which was now in verse and less prominently religious in nature.

# An Overview of Old English Literature

# The *Dictionary of Old English* Corpus

3060 “Texts”

Table 1: **DOEC statistics** for the 2009 release

A	Poetry	177,480 words	6%
B	Prose	2,128,781 words	70%
C	Glosses	699,606 words	23%
D	Glossaries	26,598 words	.88%
E	Runes	346 words	.01%
F	Inscriptions	331 words	.01%
Total		3,033,142 words	
Incl. foreign words		3,791,645 words	

# Corpus Emphases

Table 2: Standout works, authors, and genres by various metrics

Metric	Works/Authors/Genres
Most copies	Psalters (1 copy) and psalter glosses (14* copies); <i>Ælfric's Grammar</i> (14 copies)
Most productive named author	Ælfric (c. 512,500 OE words in DOEC, or c. 17 percent of the total corpus)
Most populous genres	Charters (mostly bounds; c. 1100 items in DOEC); Homilies (c. 125 works of Ælfric's in multiple copies, plus c. 350 copies of non-Ælfric texts); Saints' lives
Longest works	<i>Ecclesiastical History of the English People</i> (80,521 OE words); <i>Pastoral Care</i> (67,835 OE words); <i>History of the World Against the Pagans</i> (51,110 OE words); <i>E Chronicle</i> (47,166 OE words); <i>Consolation of Philosophy</i> (47,155 OE words [prose])

# Poetry: The Four “Poetic Codices”

Table 3: The poetic codices

Manuscript	Poetic Content Type	Poetic Works
MS Junius 11	“Biblical” poetry	<i>Genesis A, Genesis B, Exodus, Daniel, Christ and Satan</i>
Exeter Book	Riddles; lyrics; saints’ lives; religious allegory; wisdom poetry; religious poetry; misc.	c. 95 riddles and e.g. <i>Christ, Guthlac, The Wanderer, The Seafarer, Soul and Body II, The Lord’s Prayer I, The Whale, The Panther, The Wife’s Lament, The Ruin</i>
Vercelli Book	Homilies; religious poetry	<i>Andreas, The Fates of the Apostles, Soul and Body I, The Dream of the Rood, Elene, and a homiletic fragment</i>
Nowell Codex	? “Monster” poetry	<i>Beowulf, Judith</i>

# Poetry: Outside the Four “Poetic Codices”

Table 4: Poetry outside the poetic codices

MS Context	Content Type	Titles
Chronicles	Historiography	<i>The Battle of Brunanburh, The Capture of the Five Boroughs, The Coronation of King Edgar, The Death of King Edgar, The Death of Prince Alfred, The Death of King Edward the Confessor</i>
Bede	Praise/creation	<i>Cædmon's Hymn</i>
Inscriptions	Religious poetry	<i>The Dream of the Rood</i>
Fragments	Epic; historiography; wisdom poetry	<i>The Battle of Maldon, Waldere, The Finnsburgh Fragment, Solomon and Saturn</i>

## Poetry by Genre

- ▶ Biblical paraphrase: *Genesis A*, (*Genesis B*), *Exodus*, *Daniel*, *Azarias*, Psalms 51–150, *Judith*, *Christ and Satan* pt 3)
- ▶ Biblically inspired and religious narrative: *Genesis B*, *Christ*, *Christ and Satan*, *Dream of the Rood*, *Fates of the Apostles*, *Judgement Day I, II*, *Soul and Body*
- ▶ Saints' lives: *Andreas*, *Elene*, *Guthlac*, *Juliana*
- ▶ Religious allegory: *The Phoenix*, *The Whale*, *The Panther*
- ▶ Devotional: Psalms 51–150, hymns, and prayers (*Christ*, *The Descent into Hell*)
- ▶ Heroic: *Beowulf*, *Finnesburgh Fragment*, *Waldere*, *Deor*, *Widsith*, *The Battle of Brunanburb*, *The Battle of Maldon*
- ▶ Riddles, wisdom poetry, charms
- ▶ Lyric and elegy: *The Wanderer*, *The Seafarer*, *The Wife's Lament*, *The Ruin*, *Wulf and Eadwacer*, *The Husband's Message*

## Prose by Genre

- ▶ Homilies
- ▶ Laws (secular and ecclesiastical)
- ▶ Charters and records
- ▶ Saints' lives: *Dialogues* books 1–3, *Martyrology*, *Guthlac*, *Ælfric*
- ▶ Biblical translation and paraphrase: Hexateuch, Psalms, Gospels, OT selections
- ▶ Learning: *Soliloquies*, *Boethius*, computistics, *Ælfric's Grammar*
- ▶ Historiography: Orosius, Bede, *Anglo-Saxon Chronicle*
- ▶ Precepts: *Pastoral Care*, *Benedictine Rule*, *Rule of Chrodegang*, *Letter to the Monks of Eynsham*
- ▶ Liturgy and catechesis: directions, prayers, creeds
- ▶ Medical texts, recipes, charms, prognostics
- ▶ Dialogues: *Solomon and Saturn*, *Adrian and Ritheus*
- ▶ Marvels: *Wonders of the East*, *Letter of Alexander to Aristotle*
- ▶ Romance: *Apollonius of Tyre*
- ▶ Letters, tracts, and admonitions
- ▶ Notes and scribbles

## Bibliography

Healey, Antonette diPaolo, ed. “Dictionary of Old English Web Corpus.” Toronto: Dictionary of Old English Project, 2009. <https://tapor.library.utoronto.ca/doecorpus/>.

Johnson, Kyle P., Patrick J. Burns, John Stewart, Todd Cook, Clément Besnier, and William J. B. Mattingly. “The Classical Language Toolkit: An NLP Framework for Pre-Modern Languages.” In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, 20–29. Association for Computational Linguistics, 2021. <https://doi.org/10.18653/v1/2021.acl-demo.3>.