

# Dating hitherto undated early English texts based on text-internal criteria

Richard Zimmermann

Université de Genève

February 15, 2013

# Outline

## Early English Texts

- Training texts

- Test texts

## Text-internal criteria

- IP-headedness

- Scrambling

- Genitives

## Classification

## Results and Conclusion

- Saint's lives

- Medical texts

- Apocrypha & others

- Conclusion

# The general early English text corpus

- ▶ 1.2m words

## The general early English text corpus

- ▶ 1.2m words
- ▶ 5 periods: 850-910, 910-990, 990-1100, 1100-1200, 1200-1230

Period	Texts	# Words
850-910	Bede's Ecclesiastical History Gregory's Pastoral Care Boethius' Consolation of Philosophy Orosius' History Augustine's Soliloquies The Old English Martyrology The Letter of Alexander to Aristotle Marvels of the East Sections of the Anglo-Saxon Chronicle 9 <sup>th</sup> century Charters	c. 310,000

Table 1 : Early English period 1: 850-910

Period	Texts	# Words
910-990	Blickling Homilies Vercelli Homilies Rule of St. Benedict The Rule of Chrodegang of Metz West-Saxon Gospels Sections of the Anglo-Saxon Chronicle 10 <sup>th</sup> century Charters	c. 200,000

Table 2 : Early English period 2: 910-990

Period	Texts	# Words
990-1100	Ælfric's Catholic Homilies I & II Ælfric's Lives of Saints Ælfric's Supplemental Homilies Ælfric's Letters Ælfric's Old Testament Ælfric's De Temporibus Anni Wulfstan's Homilies Wulfstan's Institutes of Polity Apollonius of Tyre Alcuin's De Virtutibus et Vitiis Byrhtferth's Manual Sections of the Anglo-Saxon Chronicle 11 <sup>th</sup> century Charters	c. 530,000

Table 3 : Early English period 3: 990-1100

Period	Texts	# Words
1100-1200	Sermo in festis Sancti Marie Elucidarius Lambeth Homilies Trinity Homilies Vices and Virtues Sections of the Anglo-Saxon Chronicle	c. 80,000

Table 4 : Early English period 4: 1100-1200



Period	Texts	# Words
1200-1230	Ancrene Riwe Hali Meidhad Saint Katherin Saint Juliana Saint Margaret Sawles Warde	c. 100,000

Table 5 : Early English period 5: 1200-1230

## Undated early English texts

- ▶ 19 texts with uncertain or unknown dates of composition

## Undated early English texts

- ▶ 19 texts with uncertain or unknown dates of composition
- ▶ different genres: saint's lives, medical, apocrypha, other

Text	Ms. date	# Words	Periods
1. Saint Chad	s. XII <sup>1</sup>	2,659	1-2
2. Saint Christopher	s. X/XI	1,426	1-2
3. Mary of Egypt	s. XI in.	8,181	1-3
4. Saint Euphrosyne	s. XI in.	3,658	1-3
5. Saint Eustace	s. XI in.	5,271	1-3
6. Seven Sleepers	s. XI in	9,143	1-3
7. Holy Rood-Tree	s. XII <sup>2</sup>	6,920	1-3
8. James the Greater	s. XII med.	1,659	1-4
9. Saint Neot	s. XII med.	2,003	3-4
10. Saint Margaret	s. XII <sup>1</sup>	4,196	3-4
11. Saint Margaret	s. XI med.	3,661	1-3

Table 6 : Undated early English texts 1: saint's lives

Text	Ms. date	# Words	Periods
12. Herbarium	s. XI <sup>1</sup>	22,213	1-3
13. Quadrupedibus	s. XI <sup>1</sup>	4,276	1-3
14. Leechbook	s. X med.	34,727	1-2
15. Lacnunga	s. X/XI, XI <sup>1</sup>	7,099	2-3

Table 7 : Undated early English texts 2: medical texts

Text	Ms. date	# Words	Periods
16. Nicodemus	s. XI <sup>3/4</sup>	8,197	1-3
17. Vindicta	s. XI <sup>3/4</sup>	3,655	1-3
18. De Ascensione	s. XI <sup>1</sup>	1,798	1-3
19. Distichs of Cato	s. XI/XII	2,180	2-3

Table 8 : Undated early English texts 3: apocrypha and others

## Description of features

- ▶ 12 features measuring changes in early English

## Description of features

- ▶ 12 features measuring changes in early English
- ▶ all features are syntactic in nature



## Description of features

- ▶ 12 features measuring changes in early English
- ▶ all features are syntactic in nature
- ▶ value close to 0% - high probability of conservative variant

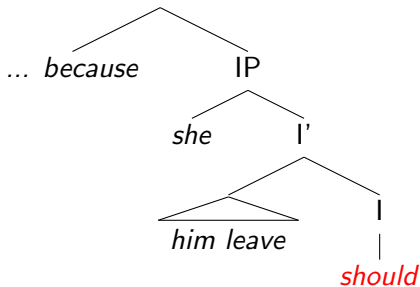
## Description of features

- ▶ 12 features measuring changes in early English
- ▶ all features are syntactic in nature
- ▶ value close to 0% - high probability of conservative variant
- ▶ value close to 100% - high probability of innovative variant

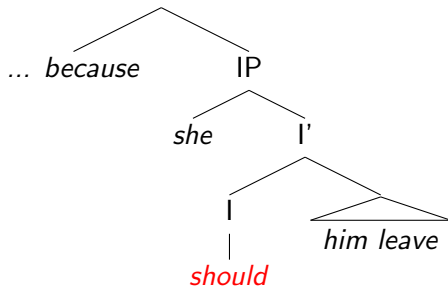
## Description of features

- ▶ 12 features measuring changes in early English
- ▶ all features are syntactic in nature
- ▶ value close to 0% - high probability of conservative variant
- ▶ value close to 100% - high probability of innovative variant
- ▶ data collected with YCOE (Taylor et al. 2003) and PPCME2 (Kroch and Taylor 2000)

## IP headedness (e.g. Pintzuk 1999)



## IP headedness (e.g. Pintzuk 1999)



## IP1 - nonfinite verbs

- (1) a. ... þæt he þæt gehat **gelæstan wolde**  
... that he that vow fulfill would  
'... that he would fulfill that vow'  
(cobede, Bede\_2:8.124.3.1170) (c. 890 A.D.)
- b. ... þæt he **wolde** swyþe lytel **drincan**  
... that he would very little drink  
'... that he would drink very little'  
(coleofri, Leof:23.27) (c. 1050 A.D.)

## IP2 - two preverbal arguments

- (2) a. ... þa [ðu] [**mæstne welan**] **hæfdest**  
... when you most wealth had  
'... when you had the greatest wealth'  
(coboeth,Bo:26.58.24.1078) (c. 890 A.D.)
- b. ... for ðan ðe [we] **habbað** [**heofenlice welan**]  
... for that that we have heavenly wealth  
'... because we have heavenly wealth'  
(ÆCHom\_II,\_38:281.63.6343) (c. 1000 A.D.)

## IP3 - non-postposing diagnostics (root clauses)

- (3) a. & þa faam of his muðe **ut eode**  
and then foam of his mouth out went  
'And then foam came out of his mouth'  
(cobede, Bede\_3:9.184.24.1845) (c. 895 A.D.)
- b. ac se biscop **eode ut** hym togeanes  
but the bishop went out him towards  
'But the bishop went out, towards him'  
(coaelhom, ÆHom\_27:43.3959) (c. 1000 A.D.)



## IP4 - non-postposing diagnostics (subordinate clauses)

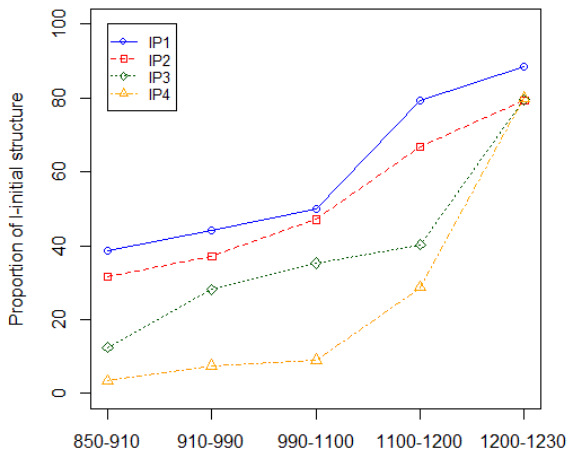
- (4) a. ... swa þa dagas **forð onsceortiað**  
... as the days forth on shorten  
'... as the days continue to grow shorter'  
(Mart\_2.1\_[Herzfeld-Kotzor]:B.13.12) (c. 875 A.D.)
- b. ... butan hi **sungon** þone lofsang **forð on**  
... unless they sang the praise-song forth on  
'... unless they continued to sing the praise'  
(coelive, ÆELS\_[Swithun]:230.4371) (c. 1000 A.D.)

## IP feature values in training texts

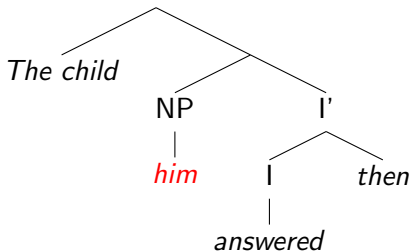
	850-910	910-990	990-1100	1100-1200	1200-1230
IP1	38.7% (762/1967)	44.2% (369/834)	50.0% (1146/2294)	79.4% (297/374)	88.4% (334/378)
IP2	31.7% (1133/3572)	37.2% (846/2277)	47.1% (2490/5290)	66.8% (342/512)	79.4% (431/543)
IP3	12.5% (144/1155)	28.2% (410/1452)	35.5% (1093/3082)	40.3% (149/370)	79.5% (435/547)
IP4	3.4% (90/2628)	7.5% (112/1496)	9.1% (308/3394)	28.7% (84/293)	80.0% (412/515)

Table 9 : IP feature values in the training material

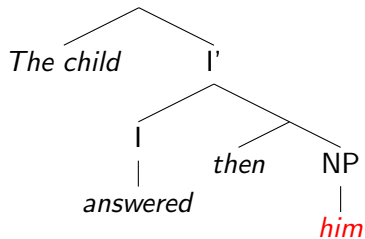
## IP feature values in training texts



## Pronominal scrambling (e.g. Bergen 2003)



## Pronominal scrambling (e.g. Bergen 2003)



## SB1 - pronominal scrambling in root clauses

- (5) a. & ic **þe wolde** eft miltsian  
and I you wanted again show mercy  
'And I wanted to be merciful to you again'  
(HomS\_11.2\_[ScraggVerc\_3]:33.411) (c. 950 A.D.)
- b. and se broðor **wolde** wurðlice **hine** bestandan  
and the brother wanted honourably him bury  
'And the brother wanted to bury him honourably'  
(coaelive,ÆLS\_[Thomas]:125.7623) (c. 1000 A.D.)

## SB2 - pronominal scrambling in subordinate clauses

- (6) a. ... þæt he sylfa **us hider** gesecean wolde  
... that he self us hither seek would  
'... that he himself would seek us here'  
(HomS\_46\_[BlHom\_11]:119.67.1519) (c. 950 A.D.)
- b. ... gif þu **geþyldlice me** gehyran wylt  
... if you patiently me hear want  
'... if you want to listen to me patiently'  
(ÆCHom\_I,\_38:514.201.7680)) (c. 1000 A.D.)

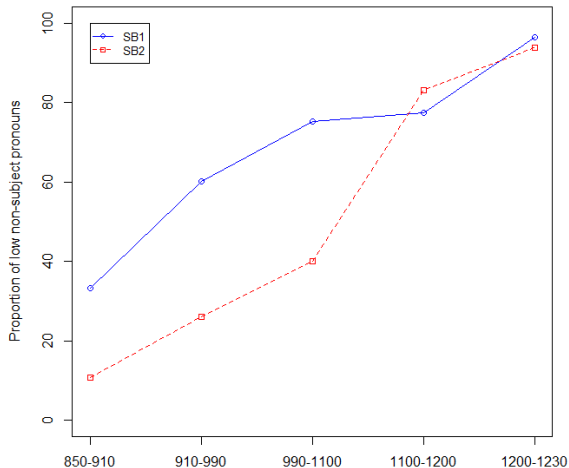
## SB feature values in training texts

	850-910	910-990	990-1100	1100-1200	1200-1230
SB1	33.2% (187/563)	60.1% (438/729)	75.3% (1138/1512)	77.3% (92/119)	96.5% (381/395)
SB2	10.8% (137/1270)	26.1% (141/540)	40.0% (476/1191)	83.1% (270/325)	93.8% (120/128)

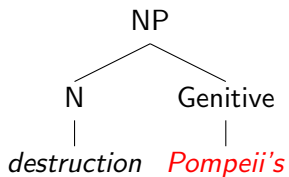
Table 10 : SB feature values in the training material



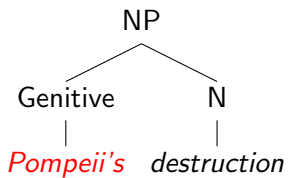
## SB feature values in training texts



## Genitive phrases (e.g. Crisma 2012)



## Genitive phrases (e.g. Crisma 2012)



## GEN - all contexts

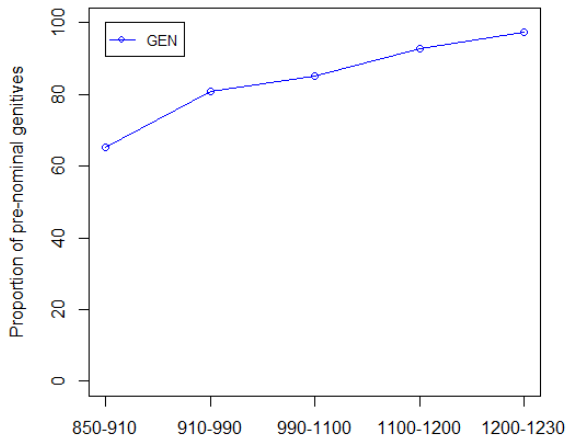
- (7) a. Her endað nu [seo æftre **froferboc Boeties**]  
here ends now the second joy-book Boethius'  
'Here now ends Boethius' second book of consolation'  
(coboeth,Bo:21.50.5.909) (c. 890 A.D.)
- b. he awrat [ða forman **Cristes boc**]  
he wrote the first Christ's book  
'He wrote the first book of Christ'  
(ÆCHom\_II,\_37:272.21.6135) (c. 1000 A.D.)

## GEN feature in training texts

	850-910	910-990	990-1100	1100-1200	1200-1230
GEN	65.2%	80.7%	85.0%	92.7%	97.3%
	(5921/9077)	(5557/5508)	(10911/12844)	(594/533)	(540/555)

Table 11 : GEN feature values in the training material

## GEN feature values in training texts



## Obtaining the data

- ▶ observation of each feature for every individual text

## Obtaining the data

- ▶ observation of each feature for every individual text
- ▶ ignore cases with too few examples



## Obtaining the data

- ▶ observation of each feature for every individual text
- ▶ ignore cases with too few examples
- ▶ spreadsheet

Early English Texts  
Text-internal criteria  
**Classification**  
Results and Conclusion

# Collection of data for each individual text

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	IP1	IP2	IP3	IP4	VP1	VP2	GEN	SB1	SB2	REL	V-C	SUB	Period
2	0.0531401	0.02195122	0.26337449	0.14668367	0.00409836	0.17687075	0.58762012	0.24324324	0.14948454	0.41368078	0.31677704	0.67816092	1
3	0.14765101	0.03944773	0.48493976	0.46899225	0.0295082	0.28494624	0.57964602	0.58823529	0.08376963	0.73245033	0.48026949	0.79518072	1
4	0.20833333	0.02516556	0.46909091	0.36852395	0.01876676	0.43870968	0.50426829	0.42424242	0.07219251	0.73051225	0.75901876	0.71502591	1
5	0.09502262	0.03598972	0.23389831	0.23899371	0.02118644	0.12371134	0.86298569	0.21637427	0.09187279	0.69411765	0.6475	0.83185841	1
6	0.02564103	0.09375	0.35135135	0.15714286	0	0.23529412	0.47311828	0.2	0.21875	0.66666667	0.26351351	0.8	1
7	0.09782609	0.0261194	0.41614907	0.39175258	0.03278689	0.31034483	0.70192308	0.5	0.11	0.67787115	0.44357977	0.64285714	1
8	0.18181818	0.05847953	0.81818182	0.52189751	0	0.44615385	0.82705611	0.49295775	0.25	0.43678161	0.45730028	0.85185185	1
9	0.14285714			0.15384615			0.85185185			0.46031746	0.63157895		1
10	0.10344828	0.05555556		0.22727273	0		0.91701245	0.18181818	0.16666667	0.54411765		0.85714286	1
11	0.1025641	0.05	0.20833333	0.23684211	0.11111111	0.16666667	0.71428571	0.35714286	0.16666667	0.73684211	0.25925926		1
12	0.16901408	0	0.21	0.0941704	0.01298701	0.25806452	0.76651982	0.41643415	0.03773585	0.87654321	0.54814815	0.71875	2
13	0.06542056	0.04347826	0.41441441	0.33823529	0.00645161	0.27472527	0.7498033	0.22916667	0.14166667	0.64793388	0.59221658	0.8125	2
14	0.09583333	0.02424242	0.3778626	0.30846774	0.02192982	0.216	0.78350515	0.19834711	0.16216216	0.72630174	0.59703947	0.76041667	2
15	0.39446367	0.11869436	0.68627451	0.54415584	0.03684211	0.45348837	0.90888031	0.81818182	0.50920245	0.8237774	0.59752475	0.84302326	2
16	0.42424242	0.09090909	0.63076923	0.25	0.08108108	0.25	0.80515298	0.66666667	0.29545455	0.91186441	0.54166667	0.81632653	2
17	0.2	0.19047619	0.26923077	0.47222222	0	0.47368421	0.84821429	0.41666667		0.67901235		0.85714286	2
18	0.14285714	0.04761905		0.55			0.82191781		0.125	0.89655172			2
19	0	0.04347826	0.76	0.69491525	0	0.57142857	0.86231884	0.66666667		0.80392157	0.73873874	0.44444444	3
20	0.21153846	0.07792208	0.61111111	0.44444444	0	0.35714286	0.75	0.65625	0.39393939	0.78787879	0.66197183	0.92857143	3
21	0.2	0.0625	0.61290323	0.22807018	0.03703704	0.33333333	0.70351759	0.5	0.10526316	0.67857143	0.54054054	0.55555556	3
22	0.35234899	0.1013986	0.51086957	0.4989154	0.10030395	0.49157303	0.90868843	0.66666667	0.48325359	0.78810976	0.78416014	0.81725888	3
23	0.28838174	0.08163265	0.56225681	0.43918919	0.02108434	0.49606299	0.79161248	0.85992218	0.40983607	0.73619271	0.75544174	0.78508772	3
24	0.35018727	0.04952381	0.52311436	0.37654909	0.02555911	0.45736434	0.805113	0.78145695	0.35828877	0.76451613	0.75286533	0.83246073	3
25	0.3943662	0.13305613	0.46082949	0.48853616	0.11340206	0.52564103	0.84727864	0.79326923	0.56551724	0.83623418	0.83676269	0.8962963	3
26	0.34951456	0.08235294	0.47169811	0.58706468	0.11538462	0.47777778	0.89530686	0.87692308	0.43478261	0.76325088	0.78181818	0.875	3
27	0.43214286	0.09782609	0.42631579	0.66153846	0.05847953	0.60952381	0.9197262	0.82213439	0.40853659	0.86465324	0.66614665	0.8558586	3
28	0	0.25714286	0	0.66	0		0.88793103		0.16666667	0.67741935	0.76088957	0.83333333	3
29	0.30107527	0.08088235	0.3960396	0.33333333	0.03030303	0.29508197	0.90588235	0.6	0.125	0.83207547	0.68086615	0.81818182	3
30	0.05383158	0.34	0.16666667	0.1	0.0388693				0.02507583	0.68865517			3

## Naïve Bayes Classifier

- ▶ probabilistic classification technique: identify to which of a set of classes (here: one of five *periods*) a new observation (here: *undated text*) belongs, on the basis of quantifiable properties (here: *the text-internal criteria*) drawn from observations whose class membership is known (here: *the general early English text corpus*).

## Naïve Bayes Classifier

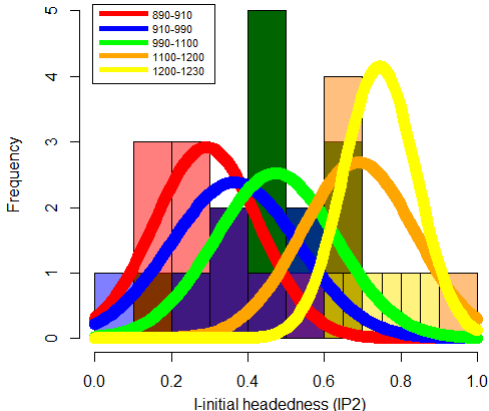
- ▶ probabilistic classification technique: identify to which of a set of classes (here: one of five *periods*) a new observation (here: *undated text*) belongs, on the basis of quantifiable properties (here: *the text-internal criteria*) drawn from observations whose class membership is known (here: *the general early English text corpus*).
- ▶ naïve: assume that feature are unrelated to each other

## Naïve Bayes Classifier

- ▶ probabilistic classification technique: identify to which of a set of classes (here: one of five *periods*) a new observation (here: *undated text*) belongs, on the basis of quantifiable properties (here: *the text-internal criteria*) drawn from observations whose class membership is known (here: *the general early English text corpus*).
- ▶ naïve: assume that feature are unrelated to each other
- ▶ many advantages: works well in complex real-world situations; robust to isolated noise; handles missing values; robust to irrelevant features; only requires small amount of training data

## Naïve Bayes Classifier - model construction

- ▶ feature values in different classes follow different probability distributions (assume feature values are distributed normally)



## Predicting new observations

- ▶ Bayes' Theorem applied to dating problem:

$$P(\textit{period}|\textit{criteria}) = \frac{P(\textit{criteria}|\textit{period}) \times P(\textit{period})}{P(\textit{criteria})}$$

## Model evaluation

- ▶ Naïve Bayes Classifier in R, package e1071

<i>predicted</i>	<i>actual</i>				
	850-910	910-990	990-1100	1100-1200	1200-1230
850-910	10	2	1	0	0
910-990	0	4	0	0	0
990-1100	0	1	13	0	0
1100-1200	0	0	0	6	0
1200-1230	0	0	0	0	6

Table 12 : Confusion matrix of training texts



## Model evaluation

- ▶ Naïve Bayes Classifier in R, package e1071
- ▶ accuracy: 39 of 43 training texts classified correctly (91%)

<i>predicted</i>	<i>actual</i>				
	850-910	910-990	990-1100	1100-1200	1200-1230
850-910	10	2	1	0	0
910-990	0	4	0	0	0
990-1100	0	1	13	0	0
1100-1200	0	0	0	6	0
1200-1230	0	0	0	0	6

Table 12 : Confusion matrix of training texts

## Results

Text	Periods	1	2	3	4	5
1. Saint Chad	1-2	1.000	0.000	0.000	0.000	0.000
2. Saint Christopher	1-2	0.735	0.260	0.005	0.000	0.000
3. Mary of Egypt	1-3	0.991	0.009	0.000	0.000	0.000
4. Saint Euphrosyne	1-3	0.015	0.978	0.007	0.000	0.000
5. Saint Eustace	1-3	0.003	0.212	0.785	0.000	0.000
6. Seven Sleepers	1-3	0.134	0.657	0.209	0.000	0.000
7. Holy Rood-Tree	1-3	1.000	0.000	0.000	0.000	0.000
8. James the Greater	1-4	0.000	0.000	0.861	0.139	0.000
9. Saint Neot	3-4	0.149	0.000	0.851	0.000	0.000
10. Saint Margaret C	3-4	0.000	0.998	0.002	0.000	0.000
11. Saint Margaret T	1-3	0.000	0.000	0.994	0.006	0.000

Table 13 : Predicted class membership of saint's lives

## Life of St. Chad

- ▶ Early date of composition predicted on basis of lexicon (Vleeskruyer 1953)

## Life of St. Chad

- ▶ Early date of composition predicted on basis of lexicon (Vleeskruyer 1953)
- ▶ Syntactic evidence confirms this:
  - mean I-initial headedness (IP2) period 1: 29.1% - Chad: 19.2%
  - mean low pronouns (SB1) period 1: 35.6% - Chad: 12.5%
  - mean pre-nominal genitives (GEN) period 1: 70.2% - Chad: 64.0%

## Life of St. Margaret

- ▶ Margaret (Corpus version) was supposedly written late (Clayton and Magennis 1994)  
evidence: inclusion of innovative words like *bread*; strong vocabulary of love

## Life of St. Margaret

- ▶ Margaret (Corpus version) was supposedly written late (Clayton and Magennis 1994)  
evidence: inclusion of innovative words like *bread*; strong vocabulary of love
- ▶ Syntactic evidence suggests that Margaret C is in fact a 10<sup>th</sup> century text

## Life of St. Margaret

- ▶ Margaret (Corpus version) was supposedly written late (Clayton and Magennis 1994)  
evidence: inclusion of innovative words like *bread*; strong vocabulary of love
- ▶ Syntactic evidence suggests that Margaret C is in fact a 10<sup>th</sup> century text
- ▶ (10) þam mannum sceal man sellan [...] beren **bread**,  
the man shall one give barley bread  
'The man should be given [...] barley bread'  
(Lch\_II\_(2)\_B21.2.1.2.2(26.1.4)) (???c. 900 A.D.)

## Results

Text	Periods	1	2	3	4	5
12. Herbarium	1-3	1.000	0.000	0.000	0.000	0.000
13. Quadrupedibus	1-3	0.000	0.000	1.000	0.000	0.000
14. Leechbook	1-2	0.149	0.196	0.655	0.000	0.000
15. Lacnunga	2-3	0.000	0.000	1.000	0.000	0.000

Table 14 : Predicted class membership of lives medical texts



## Bald's Leechbook

- ▶ Bald's Leechbook is not one uniform text but consists of three parts

## Bald's Leechbook

- ▶ Bald's Leechbook is not one uniform text but consists of three parts
- ▶ In diachronic studies, the three parts must be treated separately

# Bald's Leechbook

- ▶ Bald's Leechbook is not one uniform text but consists of three parts
- ▶ In diachronic studies, the three parts must be treated separately

Text	1	2	3	4	5
▶ 14.a Lch1	0.979	0.000	0.021	0.000	0.000
14.b Lch2	1.000	0.000	0.000	0.000	0.000
14.c Lch3	0.006	0.000	0.994	0.000	0.000

Table 15 : Predicted class membership of Leechbook 1-3

## Results

Text	Periods	1	2	3	4	5
16. Nicodemus	1-3	0.998	0.000	0.002	0.000	0.000
17. Vindicta	1-3	0.746	0.000	0.254	0.000	0.000
18. De Ascensione	1-3	0.000	0.334	0.665	0.001	0.000
19. Distichs of Cato	2-3	0.999	0.000	0.001	0.000	0.000

Table 16 : Predicted class membership of apocrypha and others

## Old English apocrypha

- ▶ The exact Latin manuscript is known on which the apocrypha are based: a collection from St. Omer, France (Cross 1996)

## Old English apocrypha

- ▶ The exact Latin manuscript is known on which the apocrypha are based: a collection from St. Omer, France (Cross 1996)
- ▶ Manuscript must have left the continent at some time before the mid 11<sup>th</sup> century

## Old English apocrypha

- ▶ The exact Latin manuscript is known on which the apocrypha are based: a collection from St. Omer, France (Cross 1996)
- ▶ Manuscript must have left the continent at some time before the mid 11<sup>th</sup> century
- ▶ Syntactic evidence suggests that both apocrypha were in fact translated early, in the 9<sup>th</sup> century

## General outlook

- ▶ Statistical classification based on syntactic features is a powerful method for chronological text ordering



## General outlook

- ▶ Statistical classification based on syntactic features is a powerful method for chronological text ordering
- ▶ Proposed early English periodisation is based on natural clusters of texts

## General outlook

- ▶ Statistical classification based on syntactic features is a powerful method for chronological text ordering
- ▶ Proposed early English periodisation is based on natural clusters of texts
- ▶ The periodisation offered might be superior to alternative classifications like the Helsinki system, which is mainly based on manuscript dates

## General outlook

- ▶ Statistical classification based on syntactic features is a powerful method for chronological text ordering
- ▶ Proposed early English periodisation is based on natural clusters of texts
- ▶ The periodisation offered might be superior to alternative classifications like the Helsinki system, which is mainly based on manuscript dates
- ▶ Future extensions: more variables, a period for late 11<sup>th</sup> century (1040-1100), separate investigation of homilies in Blickling and Vercelli collections, revise previous data on changes based on new periodisation

Thank you for your attention!

- Bergen, L. v. (2003), *Pronouns and Word Order in Old English: With Particular Reference to the Indefinite Pronoun Man*, Routledge, New York.
- Clayton, M. and Magennis, H. (1994), *The Old English Lives of St Margaret*, Vol. 9 of *Cambridge Studies in Anglo-Saxon England*, Cambridge University Press, Cambridge.
- Crisma, P. (2012), Triggering syntactic change: Inertia and local causes in the history of english genitives, in D. Jonas, J. Whitman and A. Garrett, eds, 'Grammatical Change: Origins, Nature, Outcomes', Oxford University Press, Oxford, pp. 198–216.

- Cross, James, E. (1996), *Two Old English Apocrypha and Their Manuscript Source: The Gospel of Nichodemus and The Avenging of the Saviour, with contributions by Denis Brearley, Julia Crick, Thomas Hall and Andy Orchard'*, Cambridge University Press, Cambridge.
- Kroch, A. and Taylor, A. (2000), *Penn-Helsinki Parsed Corpus of Middle English*, 2 edn, Department of Linguistics, University of Pennsylvania.
- Pintzuk, S. (1999), *Phrase Structures in Competition: Variation and Change in Old English Word Order*, Garland, New York.
- Taylor, A., Warner, A., Pintzuk, S. and Beths, F. (2003), *The York-Toronto-Helsinki Corpus of Old English Prose (YCOE)*, Oxford Text Archive.
- Vleeskruyer, R. (1953), *The Life of Saint Chad: An Old English Homily*, North-Holland, Amsterdam.