

# Variation and change in Early Middle English word order: Evidence from the Parsed Linguistic Atlas of Early Middle English

Rob Truswell, University of Edinburgh  
rob.truswell@ed.ac.uk

Multimethodological approaches to synchronic and diachronic variation, Potsdam, 25/10/19

# Research question

- ▶ There was significant syntactic variation in Middle English c.14th century.
  - ▶ (I'll show that there was even more than we realized.)
- ▶ Why did English end up with its current SVO grammar?

## More acutely

- ▶ Yang (2002): Linear Reward–Penalty algorithm for regulating grammar competition.
- ▶ Consequence: During language learning, grammars that can analyse more data are more favourably weighted.
- ▶ Corollary: ‘Once a grammar is on the rise, it is unstoppable.’ (Yang 2002: 131)
- ▶ Predicts a preference for maximally flexible grammars.
- ▶ An extremely flexible V2/V3 grammar existed in Northern Middle English (previously undescribed). And it was stopped.
- ▶ Answering the ‘why?’ question will then involve looking for models of learning that make better predictions than Yang’s in this respect.

# Roadmap

1. We made a corpus;
2. V2 in PLAEME;
3. Syntax of the Edinburgh *Cursor Mundi*;
4. Competing 14th century English grammars and their implications for the theory of grammar competition.

## Section 1

We made a corpus

## LAEME complements PPCME2

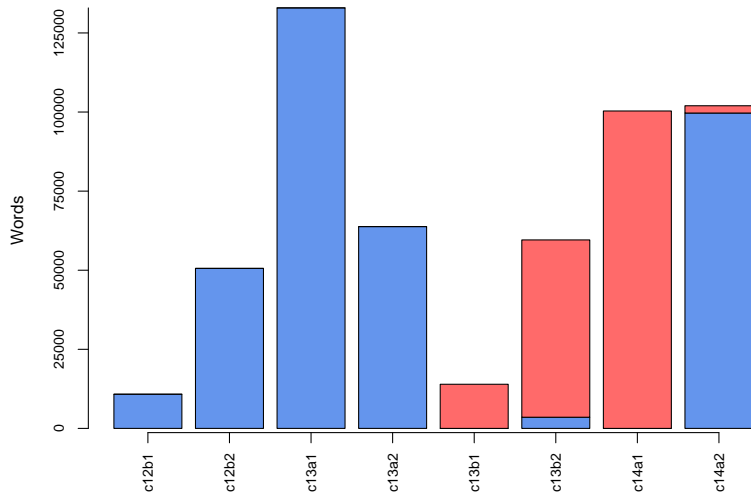
- ▶ The Penn–Helsinki Parsed Corpus of Middle English, 2nd edition (PPCME2, Kroch & Taylor 2000):
  - ▶ c.1.2m words (56 texts), covering 1150–1500;
  - ▶ POS-tagged and parsed;
  - ▶ almost exclusively prose texts (from editions);
  - ▶ virtually no data 1250–1340 (because virtually no English prose from then).
- ▶ A Linguistic Atlas of Early Middle English (LAEME, Laing 2013):
  - ▶ c.650k words (167 files) covering 1150–1325;
  - ▶ POS-tagged (very richly) and lemmatized;
  - ▶ much broader range of texts:
    - ▶ verse/prose;
    - ▶ fragmentary/whole;
    - ▶ long/short;
    - ▶ multiple versions of same text;
  - ▶ good coverage 1250–1325.

# Building PLAEME from LAEME

## Text selection

- ▶ A Parsed Linguistic Atlas of Early Middle English, Truswell et al. (2019).
- ▶ Freely downloadable at [https://github.com/rtruswell/PLAEME\\_current](https://github.com/rtruswell/PLAEME_current)
- ▶ Sample of 68 texts (172,624 words): Single version of all texts meeting the following:
  1. Manuscript is from 1250–1325;
  2. No parsed version currently exists;
  3. > 100 words.
- ▶ Where multiple versions of a text meet these criteria:
  1. Aim to balance across dialect areas;
  2. All else being equal, take the longest version.
- ▶ Small amount of text (8 files, all short) unlocalized.

# PLAEME largely fills the gap in PPCME2





## Section 2

### V2 in PLAEME

## Kroch & Taylor (1997) summary

- ▶ Southern Early ME was an IP-V2 language.
- ▶ Subject pronouns are clitics: surface higher than full NPs:
  - ▶ Some embedded V2;
  - ▶ Matrix V3 orders with subject pronouns but not with full NP subjects.
- ▶ Northern Early ME was a CP-V2 language, though subject pronouns are still clitics.
  - ▶ No embedded V2?
  - ▶ No differentiation between pronouns and full NPs w.r.t. placement of subjects.

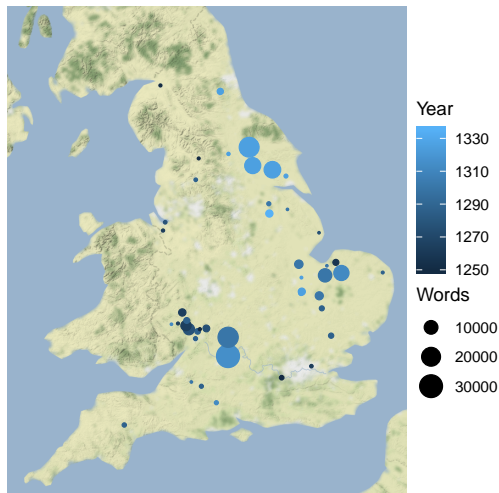
## The 'southern' pattern

- (1) Efter þe þridde fiue **3e schule** seggen [...] Kirieleyson [etc.]  
after the third five you shall say Kyrie eleison  
'After the third five, you shall say Kyrie eleison, etc.'  
(cmancriw-1-m1,l.60.193)
- (2) cheos þenne of þeos twa for þoðer **þu most** leten.  
choose then of those two for the other thou must let  
'Choose, then, between those two, because you must leave the other.'  
(cmancriw-1-m1,ll.81.978-9)
- (3) Nu **þu hauest** iseid tus.  
now thou hast said thus  
'Now you have said thus.'  
(cmhali-m1,147.276)

## The 'northern' pattern

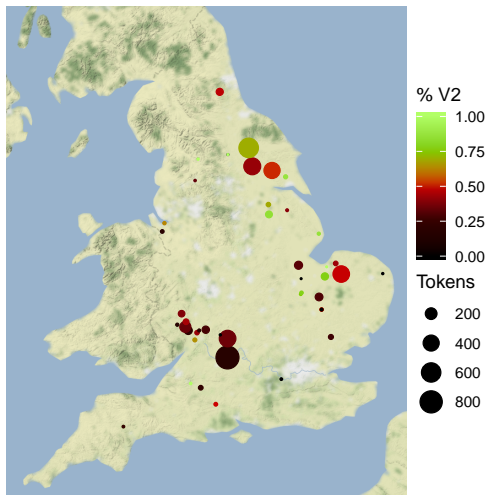
- (4) Lauerd, of me **haue** I noht,    bot þu    sende it me.  
lord      of me have I naught but thou send    it me  
'Lord I have nothing of myself unless you send it to me.'  
(cmbenrul-m3,3.60)
- (5) Mi scole **wil** i stablis    to godis seruise.  
My school will I establish to God's service  
'I will establish my school to serve God.'  
(cmbenrul-m3,4.84)
- (6) now **wil** I blinne to speke of þaim, for it ne    helpis noht  
now will I cease    to speak of them for it **NEG** helps not  
'Now I will stop speaking of them, because it doesn't help.'  
(cmbenrul-m3,5.118)

# PLAEME as dialect atlas



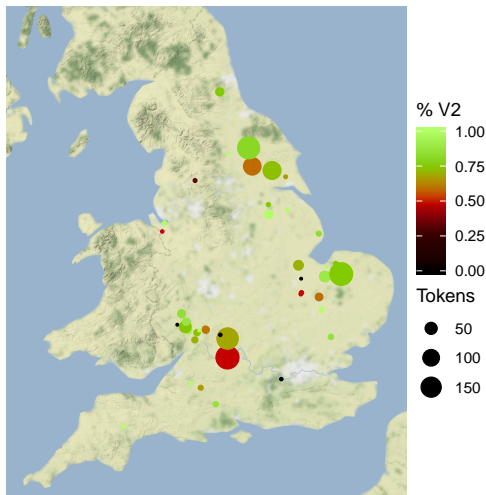
- ▶ Good representation of several broad dialect areas, though geographical coverage inevitably patchy.
- ▶ Yorkshire texts all relatively late in period, but still significantly earlier than first prose texts from the north in PPCME2.

## V2 in PLAEME



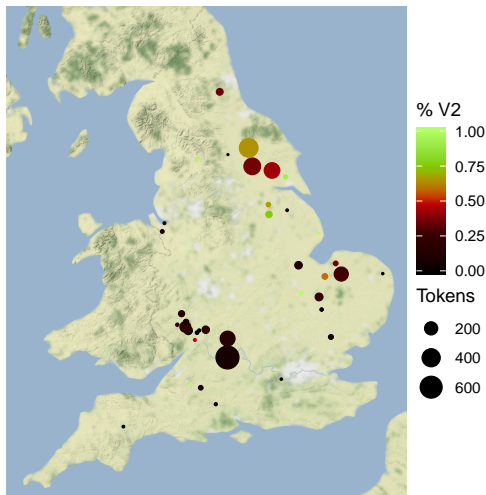
- Distribution of inversion in matrix clauses with one (or more) of the preposed elements identified by Kroch & Taylor.
- V2 concentrated in north.

## But V2 with full NPs is everywhere



- ▶ 69% of matrix clauses with preposed elements have inversion.
- ▶ No geographical pattern (no significant predictors at all).
- ▶ English c.1300 is still largely a V2 language.

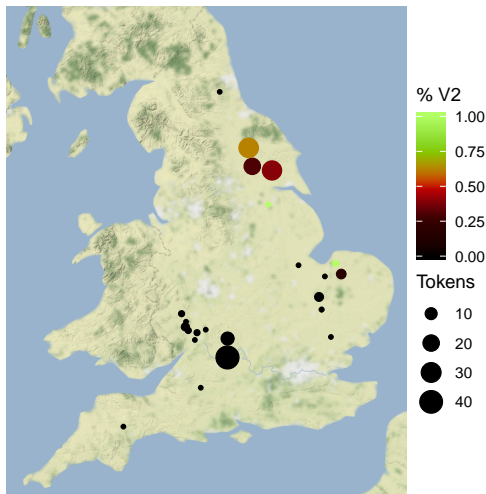
# The distinctive pronoun pattern



- ▶ Regional differences driven by inversion around pronouns.
- ▶ And inversion around pronouns in matrix clauses is a significantly northern thing.
- ▶ So Kroch & Taylor's main conclusion is supported by the PLAEME data.



## But there's more: Embedded V2 with pronouns



- ▶ Kroch & Taylor use pronoun subjects to diagnose CP-V2 vs. IP-V2.
- ▶ But inversion is also well attested in embedded clauses in some northern texts.
- ▶ Not originally taken to be a CP-V2 property.

## edincmXt is different

- ▶ Three text languages, two texts (*Cursor Mundi* in two hands, *Northern Homilies*), in one manuscript (edicmat/bt/ct).
- ▶ Main vs. subordinate clause is *not* a significant predictor of inversion.
- ▶ Can't tell in *Rule of St. Benet* because virtually no relevant contexts in subordinate clauses (only 7 vs. 343 matrix; compare edincmXt 123 embedded, 1458 matrix).

```
Formula: ifelse(Inv == "Inv", 1, 0) ~ SbjType + ClauseType +  
(1 | Filename)
```

Fixed effects:

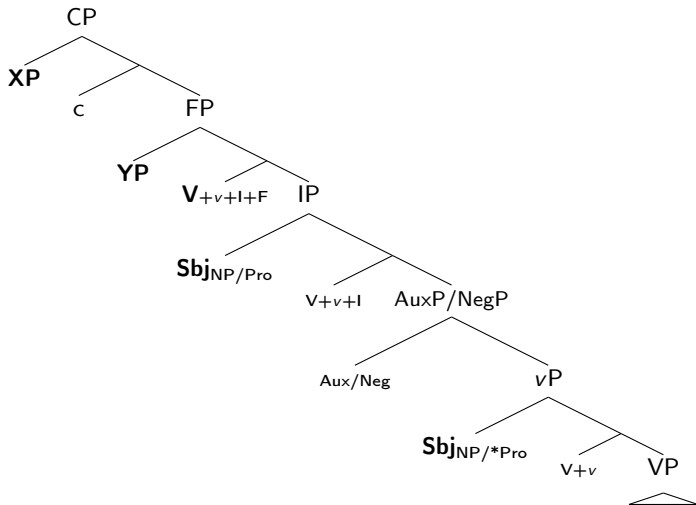
	Estimate	Std. Error	df	t value	Pr(> t )	
(Intercept)	0.69664	0.08123	2.25662	8.576	0.00912	**
SbjTypePronoun	-0.21237	0.02635	1576.11410	-8.058	1.51e-15	***
ClauseTypeSub	-0.05887	0.04457	1576.21652	-1.321	0.18675	

## Section 3

### Syntax of the Edinburgh *Cursor Mundi*

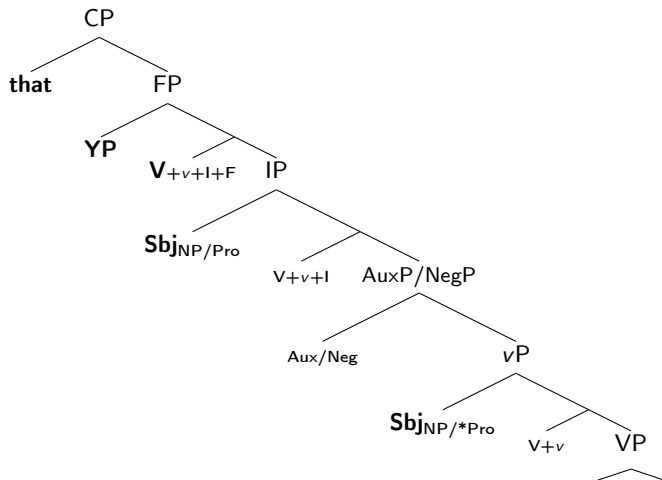
# Best guess structure

## Main clauses



# Best guess structure

## Subordinate clauses



# Predictions

## 1. Matrix V2 orders:

- (7) [Vntil hir channel] **sal** sco t(ur)ne  
until her channel shall she turn  
'It shall turn to its channel' (edincmat.75)

## 2. Matrix V3 orders (incl. XP–YP–V–Sbj):

- (8) [Sa fast] [gain oþer] **sal** tai blaw \ þat es naþing þat it  
so fast against other shall they blow that is nothing that it  
mai schaw  
may show  
'They shall blow so fast against each other that there is nothing  
that may show it.' (edincmat.108)

# Predictions

## 1. S-XP-V

- (9) [Pe cludes] [til þe heuĩ] **sal** rin  
the clouds to the heaven shall run  
'The clouds shall run to the heaven' (edincmat.137)

## 2. XP-S-V

- (10) [For drednes of þat demst(er)] \ [Pe pais] **sal** al torn ãto  
for dreadness of that deemster the peace shall all turn into  
wer  
war  
'Out of dread of that judge, the peace shall all turn into war.'  
(edincmat.52)

## Predictions

1. Postverbal full NP subjects can precede or follow participles/infinitives.

(11) [Forpi] **suld** [ilke p(re)cheour] **schau** \ Pe god þat  
therefore should each preacher show the good that  
godd hauis gert hĩ knau  
God has made him know  
'Therefore, each preacher should show the good that God has  
made him know.'  
(edincmbt.15)

(12) [Parof] **was warnid** [moyses] \ Baþe in worde ande eke ỹ dede  
thereof was warned Moses both in word and also in deed  
'Moses was warned of that, both in word and also in deed.'  
(edincmct.877)

2. Postverbal pronominal subjects must immediately follow finite element (1 counterexample out of c.600).

(13) And [þarof] **wil** [we] neu(er) blin  
and thereof will we never cease  
'And we will never cease in that.'  
(edincmat.36)



# Predictions

## 1. Embedded V2 (fronting to [Spec,FP]):

- (14) For soruĩg al dũb war þai \ [Swap̩at [a word] **miht**  
for sorrowing all dumb were they so.that a word might  
[þai] noht sai \ Na stand apõ þair fete]  
they not say nor stand upon their feet  
'For they were weeping, all dumb, so that they could not say a  
word, nor stand upon their feet.' (edincmat.931)

## 2. No general embedded V3 (XP–YP–V–S) (a few counterexamples, hard to quantify pending a better understanding of clitics, stylistic fronting, etc.):

- (15) Pan com mi cosin saint lohã \ [Pat [þan] [of welþe] **was**  
then came my cousin Saint John that then of wealth was  
[almi \ wan]]  
all.my hope  
'Then came my cousin Saint John, who was all my hope of  
wealth then.' (edincmat.1022)

## Predictions

1. No general embedded V3 (S-XP-V) (plentiful counterexamples):

(16) Noht aleue on þaĩ sal lest \ [Quēpat [þe gret] [ĩ tua]  
not a.leaf on them shall last when.that the great in two  
**sal** brest]  
shall burst  
'Not a leaf shall last on them, when the great bursts in two.'  
(edincmat.62)

2. No general embedded V3 (XP-S-V) (some counterexamples, hard to quantify as above):

(17) llworþe it es to til þe fild \ [Pat [noht] [ogain] [þe sed]  
ill.worth it is to till the field that naught again the seed  
**mai** yeld  
may yield  
'It is not worth it to till the field, when the seed may again yield  
nothing.'  
(edincmat.673)

## An aside: On the inadequacies of the best guess

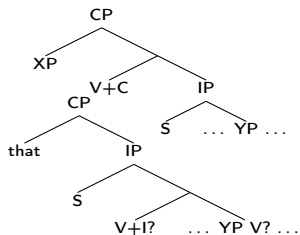
- ▶ The best guess generates a lot of orders, but it undergenerates w.r.t. embedded V3.
- ▶ This isn't an argument for anything-goes: there are categorical absences.
- ▶ The failure to predict embedded  $V > 2$  orders shouldn't be taken as evidence that this is not a V2 grammar: plentiful inversion speaks against that.
- ▶ Most likely indicates that the verb does not have to move so high in embedded clauses.
- ▶ The important point for this talk is that the best guess is a **lower bound** on what we see in these texts.

## Section 4

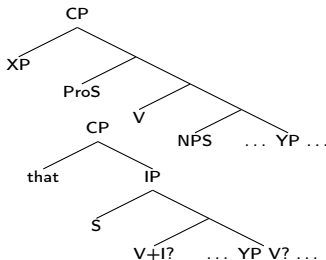
Competing 14th century English grammars and  
their implications for the theory of grammar  
competition

# At least four grammars

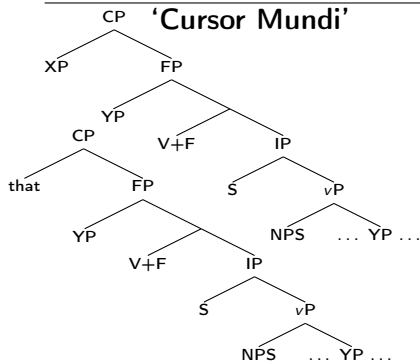
## Northern V2



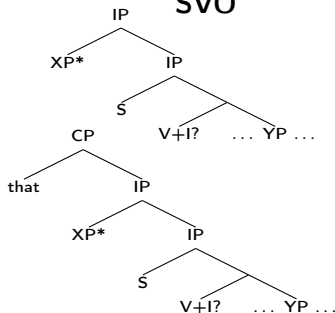
## Southern V2



## 'Cursor Mundi'



## SVO



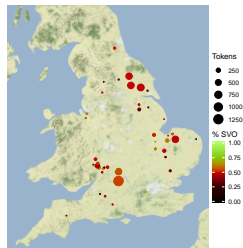
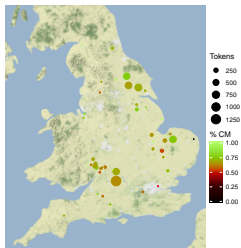
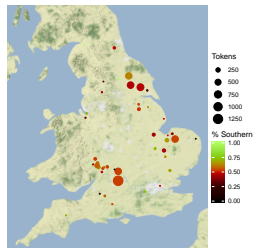
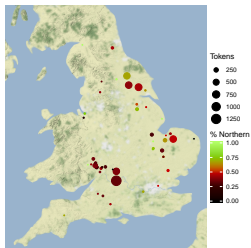
# The grammars make different predictions

	Full NP, matrix					Full NP, embedded			
	Northern	Southern	CM	SVO		Northern	Southern	CM	SVO
XP YP S V	*	*	*	✓	XP YP S V	*	*	*	✓
XP YP V S	*	*	✓	*	XP YP V S	*	*	*	*
XP V YP S	*	*	✓	*	XP V YP S	*	*	✓	*
V XP YP S	*	*	*	*	V XP YP S	*	*	*	*
XP S YP V	*	*	*	✓	XP S YP V	*	*	*	✓
<b>XP S V YP</b>	*	*	✓	✓	<b>XP S V YP</b>	*	*	*	✓
<b>XP V S YP</b>	✓	✓	✓	*	<b>XP V S YP</b>	*	*	✓	*
V XP S YP	*	*	*	*	V XP S YP	*	*	*	*
S XP YP V	*	*	*	✓	S XP YP V	✓	✓	*	✓
S XP V YP	*	*	✓	✓	S XP V YP	*	*	*	✓
S V XP YP	✓	✓	✓	✓	S V XP YP	✓	✓	✓	✓
V S XP YP	*	*	*	*	V S XP YP	*	*	*	*

	Pronoun, matrix					Pronoun, embedded			
	Northern	Southern	CM	SVO		Northern	Southern	CM	SVO
XP YP S V	*	*	*	✓	XP YP S V	*	*	*	✓
XP YP V S	*	*	✓	*	XP YP V S	*	*	*	*
XP V YP S	*	*	*	*	XP V YP S	*	*	*	*
V XP YP S	*	*	*	*	V XP YP S	*	*	*	*
XP S YP V	*	*	*	✓	XP S YP V	*	*	*	✓
<b>XP S V YP</b>	*	✓	✓	✓	<b>XP S V YP</b>	*	✓	*	✓
<b>XP V S YP</b>	✓	*	✓	*	<b>XP V S YP</b>	*	*	✓	*
V XP S YP	*	*	*	*	V XP S YP	*	*	*	*
S XP YP V	*	*	*	✓	S XP YP V	✓	✓	*	✓
S XP V YP	*	*	✓	✓	S XP V YP	*	*	*	✓
S V XP YP	✓	✓	✓	✓	S V XP YP	✓	✓	✓	✓
V S XP YP	*	*	*	*	V S XP YP	*	*	*	*

# Grammar success: All clauses



## Comments

- ▶ No grammar is close to analysing 100% of the clauses in even a single major text. 1,505 clauses in the dataset aren't analysed by any of the four grammars. This shows that I'm not working with the right grammars.
- ▶ The CM grammar clearly is able to analyse the greatest proportion of data, more or less across the board.
- ▶ This is despite the fact that the CM grammar actually *undergenerates* w.r.t. robustly attested word orders in edincmXt.
- ▶ The northern V2 grammar from Kroch & Taylor (1997) is the least successful grammar (particularly in the south).
- ▶ The other two grammars do roughly as well as each other, but without much overlap. Any regional patterns don't look very robust.
- ▶ In fact, the patterns of overlap between the grammars are *really* complex.



## Overlaps

Northern	Southern	CM	SVO	# Clauses
Y	Y	Y	Y	1641
Y	Y	Y	N	715
N	Y	Y	Y	698
N	N	Y	Y	434
N	N	N	Y	410
N	N	Y	N	372
Y	N	Y	N	296
N	Y	Y	N	216
Y	Y	N	N	84
Y	Y	N	Y	36
N	Y	N	Y	27
N	Y	N	N	7
Y	N	Y	Y	0
Y	N	N	Y	0
Y	N	N	N	0

## So why don't we use the CM grammar today?

- ▶ The CM grammar can analyse the largest quantity of data, more or less in any text.
- ▶ That should mean that, on Yang's approach, it is unstoppable.
- ▶ So how did it stop?

### Option 1: Unambiguous positive evidence

- ▶ The SVO is uniquely able to analyse more sentence tokens than the CM grammar in this dataset.
- ▶ This could be an indication that the unambiguous positive evidence for this grammar was greater than for the others.

### Option 2: Stochastic negative evidence

- ▶ The CM grammar is the most flexible w.r.t. word order.
- ▶ This means that it typically permits several alternatives to an attested order.
- ▶ So any data observed is relatively low likelihood on the CM grammar.
- ▶ This doesn't feature in Yang's algorithm, but is a crucial part of Bayesian approaches.

# Summary

- ▶ The grammar that generated edincmXt was more flexible than southern ME and than later ME grammars.
- ▶ It was flexible enough to analyse sentences from all over England.
- ▶ It could potentially have spread all over England.
- ▶ But it didn't.
- ▶ Models of learning and change should be able to predict the failure of very flexible grammars.
- ▶ Two possible directions for revising Yang's model in this respect are to take account of unambiguous positive evidence, or stochastic negative evidence.

# Acknowledgements

- ▶ Construction of PLAEME was supported by a British Academy/Leverhulme Small Research Grant (Truswell: PI; Alcorn and Wallenberg: Collaborators; Donaldson: RA).
- ▶ Thanks to Aaron Ecat, Susan Pintzuk, Beatrice Santorini, and Ann Taylor for support during the parsing process.
- ▶ Thanks also to Aaron for maintaining Annotald.
- ▶ Thanks to Caroline Heycock and Joel Wallenberg for discussion of the ideas in this talk.

## References

- Kroch, A. & Taylor, A. (1997). Verb movement in Old and Middle English: Dialect variation and language contact. In A. van Kemenade & N. Vincent (Eds.), *Parameters of Morphosyntactic Change* (pp. 297–325). Cambridge: Cambridge University Press.
- Kroch, A. & Taylor, A. (2000). Penn-Helsinki parsed corpus of Middle English (2nd edition).
- Laing, M. (2013). A Linguistic Atlas of Early Middle English, 1150–1325. Version 3.2,  
<http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html>.
- Truswell, R., Alcorn, R., Donaldson, J., & Wallenberg, J. (2019). A Parsed Linguistic Atlas of Early Middle English. In R. Alcorn, J. Kopaczyk, B. Los, & B. Molineaux (Eds.), *Historical Dialectology in the Digital Age* (pp. 19–38). Edinburgh: Edinburgh University Press.
- Yang, C. (2002). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.