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

- 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*;
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

▶ AParsed Linguistic Atlas of Early Middle English, Truswell et al. (2019).
▶ Freely downloadable at 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 ʒe schule seggen […] Kirieleysun [etc.]
After the third five you shall say Kyrie eleison
‘After the third five, you shall say Kyrie eleison, etc.’

(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.’

(3) Nu þu hauest iseid tus.
now thou hast said thus
‘Now you have said thus.’
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.
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 \textit{Cursor Mundi}
Best guess structure

Main clauses
Best guess structure

Subordinate clauses

```
CP
  /\that
 /\FP
 |   YP
 |   /\V_+v+I+F
 |   |   IP
 |   |   /\Sbj_{NP/Pro}
 |   |   |   /\vP
 |   |   |   |   \Aux/Neg
 |   |   V_+v+I
 |   |   \AuxP/NegP
 |   \V_+v+I
 |   \Sbj_{NP/*Pro}
 |   |   \vP
 |   |   \V_+v

Sbj_{NP/Pro}
```
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’

(8) [Sa fast] [gain o}ber] sal tai blaw \ \Pat es naþing \Pat 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.’


(8) [Sa fast] [gain o}ber] sal tai blaw \ \Pat es naþing \Pat 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.’
Predictions

1. S–XP–V

(9) [Þe cloudes] [til þe heuȝ] 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)] \ [Þe pais] sal al torn ĭto
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)  [Forþi]  suld  [ilke  p(re)cheour]  schau  \  Þe  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)  [Þarof]  was  warnid  [moyses]  \  Baþe  in  worde  ande  eke  ði  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 soruǐg al dūb war þai \ [Swaþat [a word] miht
for sorrowing all dumb were they so.that a word might
[þai] noht sai \ Na stand apō 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.’

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) Þan com mi cosin saint Iohã \ [Þat [ban] [of welþe] was
then came my cousin Saint John that then of wealth was
[almi \ wan]
al.my hope
‘Then came my cousin Saint John, who was all my hope of
wealth then.’
Predictions

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

(16) Noht aleue on þaĩ sal lest [Quẽþat [þe gret] [þi 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) Ilworþe it es to til þe fild [þat [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**

```
CP
  XP
  |   V+C   IP
  |   CP   S   ... YP ...
    |   that   IP
    |   S
    |   V+I?   ... YP V? ...
```

**Southern V2**

```
CP
  XP
  |   ProS   V
  |   CP   NPS   ... YP ...
    |   that   IP
    |   S
    |   V+I?   ... YP V? ...
```

**‘Cursor Mundi’**

```
CP
  XP
  |   FP
  |   YP
  |   CP
    |   V+F   IP
    |   that   FP
    |   S
    |   vP
    |   YP
    |   V+F   IP
    |   NPS   ... YP ...
```

**SVO**

```
CP
  IP
  |   XP*
  |   IP
  |   S
  |   V+I?   ... YP ...
  |   that   IP
  |   S
  |   XP*
  |   vP
  |   S
  |   NPS   ... YP ...
```
The grammars make different predictions

<table>
<thead>
<tr>
<th>Full NP, matrix</th>
<th>Northern</th>
<th>Southern</th>
<th>CM</th>
<th>SVO</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP YP S V</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>✓</td>
</tr>
<tr>
<td>XP YP V S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>XP V YP S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP YP S</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>XP S YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP S V YP</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP V S YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP S YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>S XP YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S V XP YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V S XP YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Full NP, embedded</th>
<th>Northern</th>
<th>Southern</th>
<th>CM</th>
<th>SVO</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP YP S V</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>✓</td>
</tr>
<tr>
<td>XP YP V S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>XP V YP S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP YP S</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>XP S YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP S V YP</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP V S YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP S YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>S XP YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S V XP YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V S XP YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pronoun, matrix</th>
<th>Northern</th>
<th>Southern</th>
<th>CM</th>
<th>SVO</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP YP S V</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>✓</td>
</tr>
<tr>
<td>XP YP V S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>XP V YP S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP YP S</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>XP S YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP S V YP</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP V S YP</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP S YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>S XP YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S V XP YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V S XP YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pronoun, embedded</th>
<th>Northern</th>
<th>Southern</th>
<th>CM</th>
<th>SVO</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP YP S V</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>✓</td>
</tr>
<tr>
<td>XP YP V S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>*</td>
</tr>
<tr>
<td>XP V YP S</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP YP S</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>XP S YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP S V YP</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XP V S YP</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V XP S YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>S XP YP V</td>
<td>*</td>
<td>*</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S V XP YP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>V S XP YP</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Northern</th>
<th>Southern</th>
<th>CM</th>
<th>SVO</th>
<th># Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>1641</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>715</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>698</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>434</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>410</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>372</td>
</tr>
<tr>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>296</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>216</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>84</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>36</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>27</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>7</td>
</tr>
<tr>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>0</td>
</tr>
<tr>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>0</td>
</tr>
<tr>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>0</td>
</tr>
</tbody>
</table>
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.
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 Ecay, 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


