

# Using Parsed Metrical Corpora to Investigate the Prosody-Syntax Interface

---

Raquel Montero Estebanz  
Konstanz University

Joel Wallenberg  
Newcastle University

Diachronic Generative Syntax 22  
May 2021

## Research questions

---

1. Did Heavy NP-shift in Early Yiddish create a phonological boundary between the DP and the nonfinite lexical verb?
  2. How were scrambled objects prosodically constrained in Early Middle English?
  3. Can we use poetry as Pintzuk and Kroch (1989) did in Old English to study prosodic constraints more generally?
- Answer: Yes, this methodology can be extended to study other language varieties (Yiddish and Middle English) and syntactic phenomena (scrambling).

### Generalisation:

Line breaks in poetry coincide to a high degree with the edges of phonological phrases (p-phrases).

◇ *In situ* objects: ( $\phi$ [<sub>VP</sub> O V ])

- (1) wel bid þam þe mot | æfter deaddæge / drihten secean  
well is that who may after death-day Lord seek  
'well is it with him who after his death-day may seek the Lord'  
(cobeowulf, 8.183.150 in Taylor (2008:127))

◇ Heavy NP-Shifted objects: ( $\phi$ [<sub>VP</sub>t<sub>i</sub> V ])( $\phi$ [<sub>DP</sub><sub>i</sub> O ])

- (2) þeah þe he ne meahte / on mere drifan | hringedstefnan  
though that he NEG might on sea drive ring-prowed-ship  
'though he might not steer on the sea, the ring prowed ship'  
(cobeowulf, 36.1129.928 in Taylor (2008:128))

## Background: Taylor (2005)

Extending the methodology to analyse other metrical texts from the Old English Period:

- ◇ Beowulf: alliterative verse
- ◇ Ælfrics Lives of Saints (LoS): rhythmical prose
- ◇ Metres of Boethius: alliterative verse

Text	% Nonseparated VO		% Surface VO	
Beowulf	13.33%	↓	13.64%	↓
Metres	23.53%	↓ +	38.64%	↓ +
LoS	28.28%	↓	54.29%	↓

Objects separated from the nonfinite verb by a line break are indicative of derived word order.

1. Potential Problems when using metrical texts
2. Validation
  - 2.1 Word Order in Early Middle English and Early Yiddish
  - 2.2 Mapping of prosodic constituents in poetry
3. Heavy NP-Shift in Early Yiddish Poetry (question 1)
4. Scrambling in Early Middle English Poetry (question 2)
5. Conclusion

## Potential Problems

1. In normal speech constituents can move according to IS-parameters, following different rules of prosody than in poetry. (Youmas 1983: 68)

a. I have trávell'd much in the reálm of góld

1 2 3 4 5 6 7 8 9 10

Stress Maximum Principle violation (Halle & Keyser 1971: 169).

b.

movement

Múch have I trávell'd in the reálm of góld

1 2 3 4 5 6 7 8 9 10

But, word order is not the only means to preserve the meter, so it is not necessarily used: -e endings, multiple forms of prepositions (e.g. intill/till) etc.

a. ... leffen uppo **Crist** (ID CMORM-M1, I, 231.2449)

b. ... leffen uppo **Criste** (ID CMORM-M1, I, 224.1850)

2. Metrical edges tend to correspond with the edges of prosodic constituents but need not.

# Word Order Variation in Poetry: corpora

## 1. Early Middle English

Text	Dialect	No. words	Genre	Date
Peterborough Chronicle	North East Midlands	7333	prose	c1150
The Ormulum	North East Midlands	73576	poetry	c1200
Vices and Virtues	South East Midlands	35245	prose	a1225
Poema Morale	South East Midlands	4080	poetry	a1175
Trinity Homilies	South East Midlands	41571	prose	a1225

- ◇ PPCME2-RELEASE-4 (Kroch & Taylor 2000).
- ◇ PCMEP (Zimmerman 2014/2021).

## 2. Early Yiddish

Text	Dialect	No. words	Genre	Date
Court Testimony	West Yiddish	2405	prose	1400-1490
תּוֹרַת בּוֹבּוֹ-בּוּכּוֹ (Bovo-Bukh)	West Yiddish	6363	poetry	1507
Court Testimony	West Yiddish	949	prose	1540-1589

- ◇ The Penn Yiddish Corpus (Santorini 1997/2008).

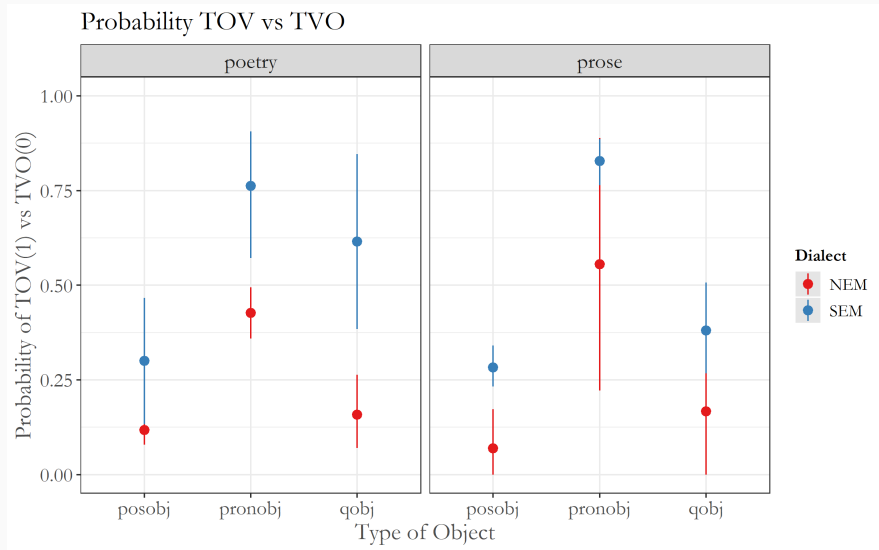
Do **dialect** (North or South East Midlands), **type of object** (positive, quantified or pronoun) and **genre** (prose or poetry) have an effect on the **position of the object** (TVO vs. TOV)?

- Results:

- ◇ **Dialect**: probability of having TOV increases when the dialect is SEM ( $p < 0.05$ ). Expected: south is more conservative.
- ◇ **Type of object**: probability of being TOV increases when the object is a pronoun, or a quantified object.  
Expected: EME quantified objects could scramble from VO order unlike positive objects (Pintzuk and Taylor 2006: 263); and pronouns could be clitics (van Kemenade 1987).
- ◇ **Genre**: no significant effect on the position of the object ( $p = 0.72$ ).



# Word Order in EME

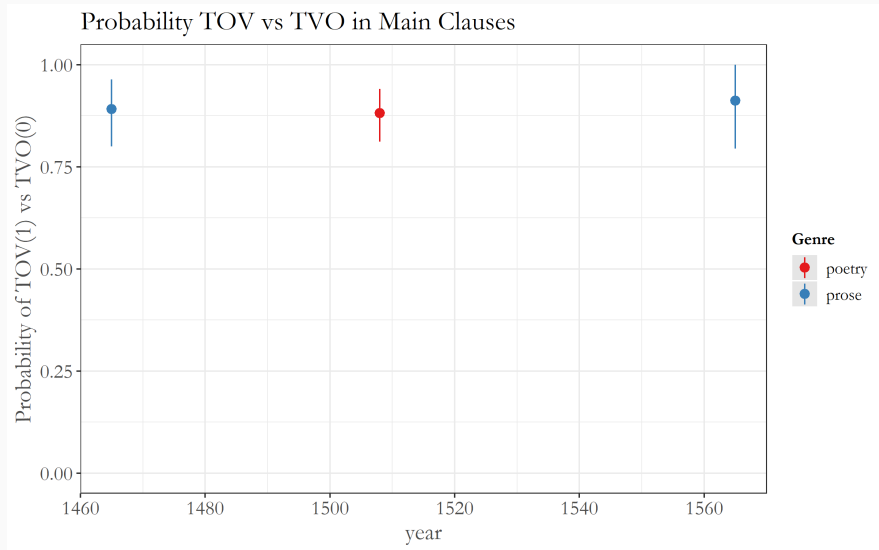


Do **genre** (prose or poetry), and **time period** have an effect on the **position of the object** (TVO vs. TOV)?

- Results:

- ◇ **Time period**: no effect on the proportions of TOV and TVO ( $p=0.752$ ).  
Expected: no attested change in headness of VPs (OV), and proportion of Heavy NP Shift is constant (Wallenberg 2009).
- ◇ **Genre**: Poetry is not significantly different from prose ( $p=0.676$ ).

# Word Order in Early Yiddish Main Clauses



# The Mapping of Prosodic Constituents

---

## Ormulum:

- Matrix clauses not beginning on a metrical edge: 1,8% (46/2441)
- Subordinated clauses not beginning on a metrical edge: 38% (325/849).
- Appositives which do not start or end on a metrical edge: 10% (33/319)
- DPs and PPs are rarely interrupted. DPs, 5 instances; PPs, 6 instances

## Bovo-Bukh:

- Matrix clauses not beginning on a metrical edge: 4% (25/653)
- Subordinated clauses not beginning on a metrical edge: 18% (26/142).
- DPs and PPs are rarely interrupted by a metrical edge.

There is a tendency to map prosodic constituents onto the metrics of the poem.

ײַב- אַװײַ (Bovo-bukh):

T-final (Santorini 1989) and OV VPs (Wallenberg 2009).

Unambiguous cases of HNPS:

1. T - V - O
2. V - T - O
3. NEG - V<sub>finite</sub> - O
4. RP - V<sub>finite</sub> - O
5. V<sub>finite</sub> - O (subordinated clauses with no V2)

Unambiguous cases of *in situ* object:

1. T - O - V (main and subordinated clauses)
2. O - V - T (subordinated clauses)

- Difference between *in situ* and HNPS:
  - ◇ HNPS: 14% (3/21) interruption by a linebreak.
  - ◇ *In situ*: 0% (0/120) interruption by a linebreak.
  - ◇ Fisher's exact test: groups are significantly different ( $p=0.0029$ ).
- Difference between Early Yiddish and Old English:
  - ◇ Both Old English and Early Yiddish had predominantly OV VPs.
  - ◇ Beowulf: 13.33% nonseparated VO (Taylor 2005); Bovo-Bukh: 85.7% nonseparated VO.
  - ◇ Difference in type of poetry:
    - Bovo-bukh 5-15 syllables per line (Harshav 2014: 179).
      1. *Vilstu es yo bobn* (Bovo-Bukh, 641.8)
      2. *do ruft men zi ale beyd zi zoltn kumen cu dem esn* (Bovo-Bukh, 201.8)
    - Beowulf 4-7 syllables per half line (Noel, 2010).

## Generalisation:

Line breaks in poetry coincide to a high degree with the edges of phonological phrases (p-phrases).

- End-base approach: (Selkirk 1986, Truckenbrodt 1999)  
Boundary of p-phrase is inserted at the right or left boundary of a filled lexical maximal projection (XP) in the syntax.
- Scrambling: (Haider 2006)  
*In situ* objects are VP internal, whereas scrambled objects are VP external.
  - ◇ *In situ* objects:  $(\phi[_{VP} O V ])$
  - ◇ Scrambled:  $(\phi[_{DP_i} O ])(\phi[_{VP} t_i V ])$
- Is there evidence from poetry that scrambled objects create a p-phrase without the lexical verb?

# Scrambling of Positive Objects in the Ormulum

---

Details: T-O-V vs T-O-Adv-V

Results: 78% interruption for scrambled vs. 13% for *in situ* OV objects.

## Analysis I:

The poems is counting syllables.

Prediction: shorter objects and adverbs make interruption less likely.

... Obj (|) Adv Verb ...  
↔            ↔

## Analysis II:

The poem is taking into account the prosody of the language.

Prediction: only the presence/absence of the adv should have an effect, because it is an scrambling diagnostic.



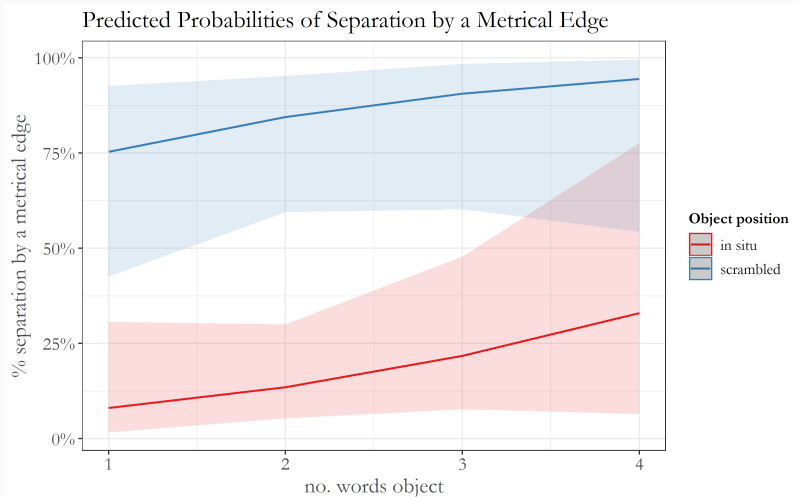
## 1. T-(Adv)-O-V vs. T-O-Adv-V:

- ◇ Dependent variable: interruption of the object from V by a metrical edge.
- ◇ Independent variable: object's length and object's position (*in situ* or scrambled).

## 2. T-Adv-O-V vs T-O-Adv-V:

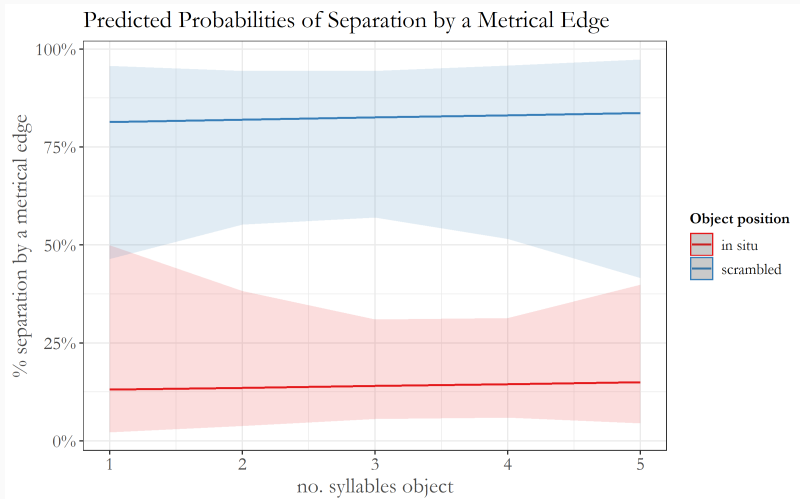
- ◇ Dependent variable: interruption of the O from V by a metrical edge.
- ◇ Independent variables: object's length, adverb's length and object's position.

# Interruption by a Metrical Edge: Results Model I



- Position object: the probability of interruption is higher when the object is scrambled ( $p < 0.05$ ).
- No. words object: no significant effect on interruption ( $p = 0.2$ ).

# Interruption by a Metrical Edge: Results Model I



- Position object: the probability of interruption is higher when the object is scrambled ( $p < 0.05$ ).
- No. syllables object: no significant effect on interruption by a metrical edge ( $p = 0.9$ ).

# Interruption by a Metrical Edge: Results Model II

1. **Logistic Regression:** convergence error, the sample size was too small for the model.

2. **Fisher's Exact Test:**

- Predictions:

Analysis I:  $\text{Length}(O + \text{Adv})_{\text{ninter}} < \text{Length}(O + \text{Adv})_{\text{inter}}$

Analysis II: length does not predict interruption.

- Grouping: Short vs Long.

2.1 þiss boc iss nemnedd Orrmulum | forrþi þatt Orrm itt wrohhte

1 2 3 4 5 6 7 8 | 1 2 3 4 5 6 7

(CMORM-M1,PREF.L1.69)

- Results: length does not predict interruption.

Short < 7 syllables  $\rightarrow p = 0.45$

Short < 6 syllables  $\rightarrow p = 0.09$

Short < 5 syllables  $\rightarrow p = 0.48$

Analysis II ✓

## Conclusion

---

1. Did Heavy NP-shift in Early Yiddish create a phonological boundary between the DP and the nonfinite lexical verb?

There is a difference between HNPS and *in situ* objects.

2. How were scrambled objects prosodically constrained in Early Middle English?

Although sparse, the data suggests scrambled objects were interrupted by a p-phrase edge from nonfinite lexical verb, which is indicated in the poem by a metrical edge.

3. Can the methodology proposed by Pintzuk and Kroch (1989) for Old English be used to study prosodic constraints more generally?

It can be extended to other languages (Yiddish and Middle English) and syntactic phenomena (scrambling).

## References

---

- Haider, H. 2006. 'Mittelfeld phenomena: Scrambling in Germanic.' In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 3. Oxford: Blackwell. 204-284.
- Halle, M., & Keyser, S. J. 1971. 'English stress.' *New York: Harper and Row*.
- Harshav, B. 2014. *Three Thousand Years of Hebrew Versification: Essays in Comparative Prosody*. Yale University Press.
- Kroch, Anthony, Ann Taylor, and Beatrice Santorini. 2000. *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, release 4 (<http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4>).
- Noel, P. A. H. 2010. 'Layers of versification in Beowulf', vol. 127, no. 2, 2010, pp. 238-260. <https://doi.org/10.1515/angl.2009.028>
- Pintzuk, S. and A. Kroch. 1989. 'The rightward movement of complements and adjuncts in the Old English of Beowulf.' *Language variation and change* 1(2): 115-143.
- Pintzuk, S. and A. Taylor. 2006. 'The Loss of OV Order in the History of English.' In A. Kemenade and L. Bettelou (eds) *The handbook of the history of English*. Oxford: Blackwell. 249-278.

## References

---

- Selkirk, E. (1986). 'On derived domains in sentence phonology.' *Phonology Yearbook* 3: 371-405.
- Santorini, B. 1989. *The generalization of the Verb-Second Constraint in the history of Yiddish*. Univ. of Pennsylvania PhD (Doctoral dissertation, dissertation).
- Santorini, B. 1997/2008. *The Penn Yiddish Corpus*. University of Pennsylvania.
- Taylor, A. 2005. 'Prosodic evidence for incipient VO order in Old English.' *English Language & Linguistics* 9(1): 139-156.
- Truckenbrodt, H. 1999. 'On the relation between syntactic phrases and phonological phrases.' *Linguistic inquiry* 30(2): 219-255.
- Wallenberg, J. C. 2009. *Antisymmetry and the conservation of c-command: Scrambling and phrase structure in synchronic and diachronic perspective*. Publicly accessible Penn Dissertations, 77.
- Youmans, G. 1983. 'Generative tests for generative meter.' *Language*, 67-92.
- Zimmermann, R. 2014/2021. *The Parsed Corpus of Middle English Poetry*. University of Geneva.