

Using Parsed Metrical Corpora to Investigate the Prosody-Syntax Interface

Going back at least to Lenerz (1977), constructions that seem serve prosodic functions have challenged a strict Principle of Phonology-Free Syntax (Zwicky 1969). For example, a number of syntactic operations can be conditioned on either the length or information structural status of constituents (e.g. Dehé 2002, Bresnan 2007), two factors that are related by their prosodic prominence. The nature of the syntax-prosody interaction is tricky to probe, and has been mostly investigated experimentally rather than in the realm of free production data that is so crucial to work in diachronic syntax. Pintzuk and Kroch (1989) pioneered using metrical texts to analyze the prosodic constraints on a syntactic operation in historical language, Heavy NP Shift in Old English (cf. also Taylor 2008). Our study uses parsed metrical texts from Old Yiddish (Santorini 1997/2008) and Middle English (Kroch and Taylor 2000) to first, replicate earlier work on HNPS and show that the same prosodic constraints hold in another Germanic variety, and secondly, to show that leftward scrambling in Germanic aligns well with certain prosodic parses and does not fit others.

Following Pintzuk and Kroch (1989), we used a set of diagnostics to confirm that to determine that uncontroversial prosodic boundaries do indeed align with metrical edges in the poetic meter (i.e. line-break or half-line breaks): clause boundaries, appositives, and vocatives. Finding that matrix clause edges and appositive constructions were interrupted less than 10% of the time by metrical boundaries, and subordinate clauses and vocatives less than 40%, we take as our starting assumption that metrical edges in the poetic meter align with prosodic boundaries, probably Intonational Phrase boundaries, where possible. We then use this information to approach the prosodic constraints on scrambling and HNPS in Middle English and Old Yiddish.

Our results show that in the Middle English *Ormulum* and the Old Yiddish *Bovo-bukh*, both scrambled and Heavy NP Shifted objects are overwhelmingly separated from the main verb by a metrical edge in the poetic meter (78% for scrambled objects and 80% for Heavy NP-shifted ones), which we would expect to align with an intonation break in the spoken language. This is illustrated in examples (1a) and (1b) for the *Ormulum* and *Bovo-bukh*, respectively (metrical edges indicated by ‘||’).

- (1) a. & tu miht ec gastlike laf || Onn operr wise jarrkenn (Scrambling)
and you might also spiritual bread || on other way prepare
‘and you should prepare God’s bread differently.’ (CMORM-M1,I,49.493)
- b. ir zult al vun im gisenkt hun || (HNPS)
you.PL shall all from him been.given have
veyn gertn un’ veld un’ vizn ||
wine garden and steppe and meadow
‘You all shall have been given by him a wine garden, a steppe and a meadow.’ (BV,35)

This constitutes evidence that scrambling in Early Middle English behaved prosodically like the scrambling of present-day German and Dutch. In present-day Germanic languages, clausal accent is assigned to an in situ object by the Nuclear Stress Rule as the deeply embedded clausal constituent. A following verb is unaccented, and so is prosodically parsed together with the object into the same Intonational Phrase. When the object has scrambled to a vP-external position, however, the verb becomes the most deeply embedded constituent in the vP. Thus, the verb is accented by the NSR, and can constitute its own intonational phrase (potentially together with vP adjuncts, etc.), and the scrambled object is parsed into an earlier intonational phrase (cf. Broekhuis 2019, *inter alia*):

- (2) a. dat Jan gisteren het **boek** gelezen heeft
that John yesterday the book read has
b. dat Jan het boek gisteren **gelezen** heeft
that John the book yesterday read has
‘that John yesterday read the book’ (Neeleman and Reinhart 1997, p. 339)

Given that scrambled and Heavy NP-shifted objects are immediately adjacent to metrical edges in the *Ormulum* in 78% and 80% of cases, on the left or right respectively, we suggest that these operations are motivated by a need to move the objects just far enough that they are not in the same intonational phrase as other *vP*-material. This phenomenon potentially has a natural explanation in phase theory, where every *vP* boundary constitutes a phase, but *only* if we accept that speakers are able to use desired prosodic outcomes to plan their phases during the syntactic computation.

The data from the *Ormulum* also shows that objects move in order to accommodate the iambic meter of the poem. For instance, in a *vP* such as (3) below, the object *man* and the verb *wurrðen* would have broken the iambic feet if the object had not been scrambled, as in (4). Out of the unambiguously 18 scrambled objects in the poem, 17 comply with the iambic feet of the poem specifically because the object has been scrambled.

(3) / x / x / x / x
 man inn hire wambe wurrðen
 man in her womb become
 ‘become human in her womb’ (CMORM-M1,I,80.704)

(4) x / x / x / / x
 inn hire wambe man wurrðen (constructed)

While fitting natural language prosody to iambic feet is at least a semi-conscious linguistic task, it is clear that the speaker does not choose to force the prosody *itself* to conform to the poetic meter (which could be done in 4 by deleting the schwa and de-accenting *man*); rather, scrambling is used to protect both the NSR and the meter. Moreover, we show that the overall syntax of the *Ormulum* is statistically nondistinct (at least object placement) from that of other texts from the same period and dialect area, and so we can be relatively confident that the text is a reasonable representative of the language spoken in the East Midlands.

In the Old Yiddish *Bovo-bukh*, the absence of half-lines in the meter makes it more difficult to detect intonation breaks, but the data still broadly show the same pattern as in the *Ormulum* (14.3% of HNPS objects occur after a line-break, compared to 0.3% for other objects).

Parsed metrical corpora are a highly important tool, not only for gaining prosodic information about dead languages, but also for investigating the syntax-prosody interface quantitatively in free-production data. The fact that poetry/song may be a human universal, and speakers habitually produce it outside of laboratory settings, and in languages that may be difficult to access in such settings, means that parsed metrical corpora are likely an under-used resource. We’d also suggest the field might combine the quantitative, comparative, and metrical dimensions in creating parallel parsed metrical corpora.

- Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? experiments with the english dative alternation. *Roots: Linguistics in search of its evidential base* 96:77–96.
- Broekhuis, Hans. 2019. The unification of object shift and object scrambling. In *The Cambridge Handbook of Germanic Linguistics*, ed. M. Putnam and Michael P. Richard. CUP.
- Dehé, Nicole. 2002. *Particle verbs in english: Syntax, information structure and intonation*, volume 59. John Benjamins Publishing.
- Kroch, Anthony S., and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English. CD-ROM. Second Edition, release 4. Size: 1.3 million words.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: TBL-Verlag Narr.
- Neeleman, Ad, and Tanya Reinhart. 1997. Scrambling and the pf interface. In *The Projection from Arguments*, ed. M. Butt and W. Gueder, 309–353. Stanford: CSLI Publications.
- Pintzuk, Susan, and Anthony S Kroch. 1989. The rightward movement of complements and adjuncts in the Old English of Beowulf. *Language Variation and Change* 1:115–143.
- Santorini, Beatrice. 1997/2008. The Penn Yiddish Corpus. University of Pennsylvania.
- Taylor, Ann. 2008. A note on the position of the verb in Old English rhythmic prose. *York papers in linguistics Series 2* 124–141.
- Zwicky, Arnold M. 1969. Phonological constraints in syntactic descriptions. *Research on Language & Social Interaction* 1:411–463.