

## The variable syntax of pronominal prepositional objects in Old English: Some statistical modeling problems

Rhona Alcorn, University of Edinburgh  
rhona.alcorn@ling.ed.ac.uk

The variable positioning of personal pronouns in Old English prose remains something of a mystery. In the role of prepositional object, for example, personal pronouns can be found in the right-of-P position, as in (1), or in a variety of left-of-P positions, e.g. (2). With very few exceptions, only the right-of-P position is available when the preposition's object is a demonstrative pronoun or nominal, allowing left-of-P placement to be interpreted as 'special' placement.

- (1) Ac þa hundas comon **to him**  
*But the dogs came to him*  
'But the dogs came to him'  
(cogregdC,GDPref\_and\_4\_[C]:34.310.6.4623)
- (2) a. and se hælend sylf of heofonum com **him to**  
*and the Saviour Himself from heaven came him to*  
'and the Saviour Himself came to him from heaven'  
(coaelive,ÆLS\_[Thomas]:13.7546)
- b. ond misenlico wilddeor **him** þær comon **to**  
*and various wild beasts him there came to*  
'and various wild beasts came to him there'  
(comart3,Mart\_5\_[Kotzor]:Ju2,A.6.887)

There is ample evidence to suggest that such special placement is not the freely available option predicted by leading analyses. Instead, it appears to be conditioned by a range of linguistic and extra-linguistic factors, making the dataset represented by (1) and (2) an ideal candidate for logistic regression. Before such an analysis can be undertaken, however, three important issues must be addressed:

1. given certain prepositions are formally identical to certain verbal prefixes and/or adverbs, how reliable is the labelling of prepositions in the YCOE?
2. given third person data suggest that left-of-P placement rarely occurs unless the pronoun has dative case (as I will show), how can this effect be controlled for first and second person data, whose dative and accusative forms are identical?
3. given there are insufficient tokens to include all factors and all factor groups in a logistic regression model, how can the 'best' model be constructed?

I will provide an answer to each of the first two questions and, time permitting, will suggest a way of approaching the problem described by question three.