

# The information structure of subject extraposition in Early New High German

Caitlin Light

University of Pennsylvania

August 23, 2011

## Introduction

- This talk investigates the information-structural characteristics of extraposed subjects in Early New High German (ENHG).
  - Using quantitative data from a parsed corpus of ENHG, I will compare the behavior of extraposed subjects to the behavior of extraposed objects, as described by Bies (1996).
  - It will be shown that in ENHG, subject extraposition shares some characteristics with object extraposition.
    - Both may be realizations of narrow focus on the extraposed DP.
  - However, subject extraposition may also be motivated by presentational focus.
    - In this sense, extraposition is one possible means to obtain default accent on a subject.
  - This shows a potential parallel between ENHG and Yiddish:
    - Prince (1989) has demonstrated that Yiddish subject extraposition is motivated by the discourse-newness of the subject.

## Outline

- 1 Introduction
- 2 Previous studies
  - Object extraposition in ENHG
  - Subject extraposition in Yiddish
- 3 Methodology
  - The corpus used
  - Terminology
  - Information-structural classification of sentences
  - Comparing non-extraposed subjects
- 4 Results
  - Overall properties of subject extraposition
  - The information structure of subject extraposition
- 5 Analysis
  - Narrow focus
  - Presentational focus
  - Extraposition and sentence accent
- 6 Conclusion and implications

## Bies (1996): object extraposition in ENHG

- Bies (1996) provides a detailed analysis of the information structure of DP extraposition, based on a corpus of examples collected from ENHG texts.
- She considers two possible motivations for extraposition:
  - Discourse newness of the DP
  - Narrow focus on the DP more generally
- Before considering discourse factors, Bies identifies external influences on DP extraposition:
  - First, definiteness of the DP has a clear effect on extraposition:

	<b>Postposed</b>	<b>Non-post.</b>	<b>Rate of post.</b>
<b>Negation</b>	0	86	0%
<b>Indefinite</b>	8	301	2.3%
<b>Other QP</b>	2	83	2.4%
<b>Non-quant.</b>	64	408	13.6%

## Bies (1996): general discourse properties of object extraposition

- Second, DP length (or ‘weight’) also strongly influences extraposition:

	<b>Postposed</b>	<b>Non-post.</b>	<b>Rate of post.</b>
<b>Conjoined</b>	13	17	43.3%
<b>Relative clause</b>	5	2	71.1%
<b>PP postmodifier</b>	25	34	42.4%
<b>Regular length</b>	82	722	10.2%
<b>Total</b>	125	775	13.9%

- Based on these observations, Bies restricts her data set to ‘regular length,’ non-quantified DPs.
- She also excludes topicalized and scrambled objects from the set of non-extraposed DPs, assuming that they represent unrelated information structural phenomena.

## Bies (1996): quantitatively evaluating information structural analyses

- Bies separates her (restricted) data set into three informational categories:
  - Discourse-new information
  - Evoked/inferred information
  - Given information
- She observes a gradient relationship between newness and extraposition:

	<b>Postposed</b>	<b>Non-postposed</b>	<b>Rate of postposing</b>
<b>Given</b>	11	100	10%
<b>Evoked/Inf.</b>	37	81	31.4%
<b>Disc.-new</b>	16	21	43.2%
<b>Total</b>	64	202	24.1%

- However, Bies asks an additional question: is this the main discourse motivation for DP extraposition, or a symptom of it?

## Bies (1996): quantitatively evaluating information structural analyses

- Bies recognizes the fact that discourse-newness of a DP may contribute to its likelihood of being narrowly focused: discourse-new, or presentational, elements often are the focus of a sentence.
  - She suggests the relationship between discourse-newness and extraposition is simply a consequence of the fact that these elements are more likely to be narrowly focused.
  - To explore this alternate hypothesis, she further classifies her sentences into (narrow) DP focus and (wide) VP focus, wherever context allowed an unambiguous classification.

	<b>DP focus</b>	<b>VP focus</b>	<b>Percent DP focus</b>
<b>Non-postposed DP</b>	19	123	13.4%
<b>Postposed DP</b>	46	4	92%

- The effect of narrow focus is stronger than that of information status.
- Bies therefore concludes that narrow focus in a general sense motivates object extraposition in ENHG.

## Prince (1989): subject extraposition in Yiddish

- One could simply assume that this is the general pattern for DP extraposition in ENHG, but another analysis is possible.
- Prince (1989) considers the information structural properties of subject extraposition in Yiddish, a closely related language, and argues that in this case, the extraposition is motivated by the discourse new status of the DP.
- This raises a question: should ENHG have one unified analysis for DP extraposition as a whole, or do subject and object extraposition behave differently?



## Prince (1989): quantitative support for the new information focus analysis

- Prince begins by considering a set of examples from a parsed corpus of Yiddish.
- She finds that in certain subordinate clause types, brand-new subjects are highly motivated to postpose:

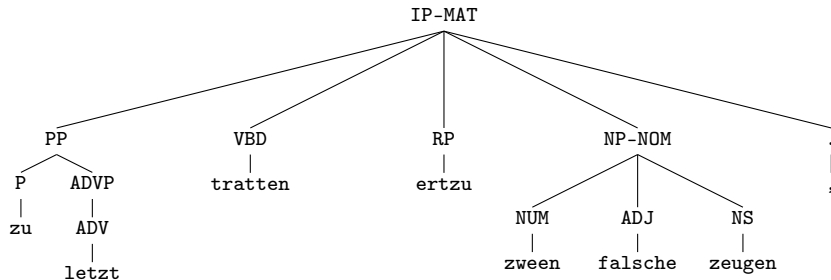
	<b>Non-postposed</b>	<b>Postposed</b>	<b>Total</b>
<b>Adverbial</b>	16 (57%)	12 (43%)	28
<b>Complement</b>	14 (30%)	32 (70%)	46
<b>Total A/C</b>	30 (41%)	44 (59%)	74
<b>Free Relative</b>	0 (0%)	7 (100%)	7
<b>Indirect Q.</b>	0 (0%)	3 (100%)	3
<b>Relative</b>	0 (0%)	2 (100%)	2
<b>Total WH</b>	0 (0%)	12 (100%)	12

## Prince (1989): syntactic support for the new information focus analysis

- She also proposes a syntactic motivation for her analysis:
  - Expletive *es* ('it') is licensed to fill Spec,TP when the subject is extracted from a free relative or indirect question (cf. Diesing, 1990; Prince, 1989). This does not occur with subject relative clauses.
    - (1) a. Ikh veys nit ver **es** iz gekumen  
I know not who ES is come  
'I don't know who came.'  
b. \*Ikh veys nit ver iz gekumen
    - (2) a. Der melamed vos iz besser far ir iz beser far mir.  
the teacher that is better for her is better for me  
'The teacher that is better for her, is better for me.'  
b. \*Der melamed vos **es** iz beser far ir iz beser far mir.
  - Prince hypothesizes that this expletive appears when the subject is extracted from a **postposed position**, leaving Spec,TP empty.
  - This explains why the expletive appears only in indirect questions and free relatives, which (unlike relative clauses) have an extracted element that is new to the discourse.

## The *Septembertestament* corpus

- The data for this study was drawn from a parsed corpus of Martin Luther's *Septembertestament* bible translation, published in 1522.
- The *Septembertestament* corpus consists of roughly 102,000 words.
  - The corpus was initially parsed by automatic methods (including Bikel, 2004), but ultimately hand-corrected by the author according to the guidelines for the Penn Historical Corpora and the York-Torono-Helsinki Parsed Corpus of Old English Prose (cf. Kroch, Santorini, and Diertani, 2004; Kroch and Taylor, 2000; Taylor et al., 2003), adapted for use for a German corpus.
- Example of a simple parsed sentence:



## Using the corpus to study information structure

- A parsed corpus of this type is not fully equipped for information-structural or discourse-related research.
  - Once relevant tokens were extracted from the corpus, they were hand-coded for additional information to make them usable for this purpose.
  - The following subsection explains the methodology used in processing examples of extraposed (and potentially extraposed) subjects.
- I will begin, however, by commenting on my use of information structural assumptions and terminology in the study.

## Theoretical assumptions and terminology

- I roughly assume a theory of information structure as in Vallduví (1992), in which every sentence is divided into a focus-ground partition, and every sentence has only one information-structural focus.
- I use the term ‘narrow focus’ to describe a DP which is, in itself, the sole focus of a clause (that is, neither part of the ground, nor part of a larger focused constituent).
  - The term ‘narrow focus’ in this sense quite general.
- I also make use of the term ‘presentational focus,’ which I take to describe the type of focus which introduces a new entity into the discourse.
  - I will claim that presentational focus may either occur narrowly on the DP introducing the new entity, or it may involve broader focus, as I will demonstrate in a later section of the presentation.
  - Other narrowly focused DPs may be interpreted as other types of focus, for example contrastive focus, which I will not discuss in detail today.

## Unambiguous examples of subject extraposition

- Because the German verb phrase is known to be right-headed (in traditional terminology), a verbal element may be used to diagnose the right edge of the verbal domain.
  - Phrases appearing to the right of such a diagnostic are said to be extraposed.
  - This excludes the finite verb in verb-second clauses, which has been raised to C.
  - Therefore, clauses with a non-finite verb or particle are treated as unambiguous environments for extraposition.
  - In addition, the finite verb in subordinate clauses with an overt complementizer may diagnose the right-edge verb position (although CP recursion may be an issue I must address in the future).
- For this study, I extracted all nominative subjects that appear in an extraposed position.
- 115 subjects were found which fit these requirements.
- All clauses were hand-coded for definiteness, syllable length, discourse status (based on Bies's classifications), and narrow DP focus on the subject vs. other focus structures (where possible).

## Comparing subjects left *in situ*

- I then proceeded to extract all clauses with subjects that were capable of extraposing, but did not.
  - These clauses must contain the same diagnostics; that is, a final non-finite verb or verbal particle.
  - Pronominal subjects were excluded, including impersonal *man* ('one'), as well as demonstrative determiners; these are all assumed to be incapable of extraposition.
- A total of 1261 non-extrapolated subjects satisfying these criteria were found.
- Each of these tokens was hand-coded for syllable length and the definiteness of the subject.
- Different sub-samples of this set were isolated for the consideration of the pragmatic and information-structural characteristics of subject extraposition, as I will discuss in a moment.

## Overall properties of subject extraposition

- The overall rate of subject extraposition was roughly 8.4%.
- This is lower than the overall rate for object position given in Bies, 13.2%.
- Additionally, as Bies showed for object extraposition, weight proved to be a strong influence on subject extraposition.
  - I deviate from Bies in measuring DP weight by syllables, rather than by modifier presence and type.
    - For extraposed subjects, the minimum weight was **2** syllables, while the maximum was **64** (due to a sequence of embedded clauses within the DP).
    - The average weight of extraposed subjects was **12.6** syllables.
    - For non-extraposed subjects, the minimum weight was **1** syllable, while the maximum was **29**.
    - The average weight of non-extraposed subjects was only **3.42** syllables.



## Adjusting the sample

- To minimize the effect of DP weight on the sample, I chose to limit my sample to subjects of 15 syllables or less.
  - This ensures that the DPs are of a weight safely below the limit found on non-extrapolated subjects, without too greatly restricting the data set.
  - The remainder of the paper will deal only with this subset of the data, unless otherwise noted.
- The adjusted sample includes 86 extrapolated subjects and 1257 non-extrapolated subjects, or extraposition at a rate of about 6.4%.
- Subject extraposition in this sample is more frequent in matrix than subordinate clauses:
  - There are 64 extrapolated and 654 non-extrapolated subjects in matrix clauses, or 8.9% extrapolated.
  - There are 20 extrapolated and 602 non-extrapolated subjects in subordinate clauses, or 3.2% extrapolated.

## Overall properties of subject extraposition

- This fact concerns the entire data set, regardless of DP weight.
- When we consider DP type, we find a striking difference between subject and object extraposition: indefinite and quantified/numeric subjects are extraposed more, and definite subjects less frequently.

	Non-extraposed	Extraposed	% Extraposed
<b>Negative</b>	47	0	0.0%
<b>QP/Numeric</b>	76	17	18.28%
<b>Indefinite</b>	98	9	8.41%
<b>Bare</b>	12	13	52.00%
<b>Free rel.</b>	13	28	68.29%
<b>Definite</b>	1015	48	4.52%
<b>Total</b>	1261	115	8.36%

- But note that according to the data, negative subjects never extrapose.
- Additionally, bare subjects extrapose more often than not.
  - It is unclear how this relates to Bies's data, as her examples suggest she may have decided to include bare DPs in here "non-quantified" category.

## Pragmatic properties of subject extraposition in ENHG

- I compared the discourse status of all extraposed subjects to non-extraposed subjects in a subset of 443 matrix clauses and 173 subordinate clauses (for purposes of examining a smaller subset of the data). In this set I exclude subjects of over 15 syllables.

	Discourse-new	Evoked/Inf.	Given
<b>Extraposed</b>	<b>33 (38.82%)</b>	26 (30.59%)	26 (30.59%)
<b>Non-extraposed</b>	75 (12.17%)	165 (26.79%)	<b>376 (61.04%)</b>

- While the majority of extraposed subjects are discourse-new, the majority of non-extraposed subjects are given.

## Narrow focus and extraposed subjects

- The 86 clauses with extraposed subjects were then compared to a randomly selected sample of 60 clauses with non-extraposed subjects, for a detailed consideration of the focus structures of these groups.
- The "other" category includes VP and broader focus, as well narrow focus on a constituent other than the subject.
- Examples with ambiguous focus structures were set aside.

	<b>Narrow S-foc.</b>	<b>Other foc.</b>	<b>% Narrow foc.</b>
<b>Extraposed</b>	36	38	48.7%
<b>Non-extraposed</b>	4	41	8.9%
<b>Total</b>	40	79	33.6%

- We can see that extraposed subjects are narrowly focused more often than non-extraposed subjects. However, whereas Bies found that 92% of extraposed objects were narrowly focused, only 48.7% of subjects are narrowly focused in my sample.
- This leads me to reject the hypothesis that subject extraposition is driven solely by narrow focus, as Bies has proposed for objects.

## Subject extraposition and narrow focus

- As the data show, there are some ways in which subject and object focus pattern similarly.
- Numerous examples can be found which involve narrow focus on an extraposed definite subject:
  - vnnnd eynem gab er funff centner, dem andern zween, dem dritten  
 and one gave he five talents the other two the third  
 eyn, eynem ydern noch seynem vermugen vnd zoch hynweg ... **vnd**  
 one one each after his ability and went away ... and  
**da tratt ertzu, der da funff centner empfangen hatte**  
 then tread forward who PART five talents received had  
 ... **Do trat auch ertzu, der do zween centner**  
 ... Then tread also forward who PART two talents  
**empfangen hatte** ...  
 received had ...

‘And he gave five talents to one, two to another, one to the third,  
 each according to his ability, and went away ... and then the man  
 who had received five talents came forward ... Then also, the one who  
 received two talents came forward ...’  
 (*Septembertestament*, Matthew 25:15–22)

## Subject extraposition and narrow focus

- In some such cases, a focus particle makes the focus structure especially clear:
  - (4) ... **das** **verfuret** **werden**, **yhn** **denn** **yrthum** **wo** **es** **muglich**  
...that misled will-be in the confusion where it possible  
**were** **auch** **die** **auserweleten**.  
would-be also the chosen  
'so that in the confusion, where possible, even the chosen will be misled.'  
(*Septembertestament*, Matthew 24:24)
- I argue that in cases such as these, subject and object extraposition do in fact have the same motivation: narrow focus on the DP.
- However, the frequencies at which quantified subjects extrapose, as well as discourse-new/presentational subjects, makes this look more like Yiddish, where the discourse status of the subject plays a more important role.
- Therefore, I will argue that subject extraposition in ENHG can also be motivated by a more specific type of focus.

## Presentational clauses and subject extraposition

- In German, the default sentence accent falls near the end of the clause, most often on the rightmost argument, immediately preceding a sentence-final verb or particle (cf. Ladd, 1996; Truckenbrodt, 2007).
  - As a result, object DPs are frequently in the appropriate position to receive default sentence accent.
  - Scrambling of other elements can place the object at the right edge of the ‘middle field,’ so that it receives default accent when necessary.
  - Extraposition of object DPs may be expected to have more specific motivations than simply to obtain default accent; Bies’s analysis seems to support this. Object DPs extrapose specifically to receive narrow focus.

## Presentational clauses and subject extraposition

- However, if the sentence accent falls on a constituent near the left edge (either topicalized or as the leftmost constituent in a Tense-final subordinate clause), it results in an obligatorily contrastive interpretation (cf. Frey, 2006).
  - We may hypothesize that extraposition of subject DPs may be usable as a general means to obtain default accent on a subject, without resulting in a contrastive interpretation.
  - There is a specific clause type that may demonstrate this: clauses with presentational focus.



## Extrapositing presentational subjects

- Indeed, a large subset of the extraposed subjects occur in presentational contexts, and best translated with the use of existential *there*.
  - (5) Aber die kinder des reychs werden außgestossen ynn die but the children of-the kingdom will-be cast-out in the außersten finsterniß, **da wirt seyn weynen vnd tzeen klappen.** outermost darkness there will be weeping and teeth gnashing  
'But the children of the kingdom will be cast out into the outermost darkness. In that place there will be wailing and gnashing of teeth.'  
(*Septembertestament*, Matthew 8:12)

## Extraposition of presentational subjects

- Other clauses have the same effect of introducing a new entity into the discourse, but do not appear to involve narrow focus on the DP.

(6) Da stund aber auff ym radt **eyn Phariseer mit**  
then stood however up in-the council a Pharisee with  
**namen Gamaliel, eyn schriftgelerter, wolgehallten fur allem**  
name Gamaliel a scholar well-held before all  
**volck**  
people

‘But then a Parisee named Gamaliel stood up in the council, a scholar, well regarded by all the people.’

(*Septembertestament*, Acts 5:35)

## Characteristics of presentational subjects

- The nature of these subjects implies that they will generally be quantified, indefinite or bare NPs. Additionally, they will generally be new entities in the discourse.
- We can also observe this effect by considering the occurrence of copular clauses in each data set: while 16 (18.6%) of extraposed subjects occur in copular clauses, only 92 (7.3%) of non-extraposed subjects do.
  - This means that copular clauses extra pose at a rate of 14.8%, while non-copular clauses extrapose at a rate of 5.67%.
  - My argument is that the link between subject extraposition and copular clauses is due to the fact that many copular clauses are presentational, and thus favor extraposition of the subject above other clause types.
- This analysis provides an explanation for the fact that quantified objects extrapose less often, and quantified subjects more often, than definites.

## A theory of assigning sentence accent

- Ladd (1996) discusses a metrical account of sentence accent.
  - He observes that the accent patterns of the following two sentences may differ, when both are interpreted with broad (sentential) focus:

(7) a. JOHNSON died.  
b. Former president Johnson unexpectedly DIED today.
  - He proposes that this can be explained by the fact that the shorter utterance may consist of only one intermediate intonational phrase, and within the intermediate phrase the primery accent falls on the subject.
  - However, once the utterance is as long as (7b), the subject and predicate may not form a single intonational phrase, and must be split into two intermediate phrases.
  - These two phrases have a weak-strong relation, meaning that the primary sentence accent falls on the strongest accent in the second intonational phrase.
  - “The heavier a constituent is, the more likely it is to constitute its own intermediate phrase.” (Ladd, 1996)

## Sentence accent and the extraposed subject

- Wallenberg (p.c.) proposes that in English Heavy NP-Shift (HNPS), a "Heavy NP" moves rightward past any material on its right in order to constitute its own (rightmost) intermediate phrase, and thus obtain primary accent.
- I tentatively propose that the same can be true of subject extraposition in ENHG.
  - In order to receive primary accent while allowing a presentational (and non-contrastive) focus interpretation, the subject may move to the right edge.
  - By extraposing, the subject may form its own intermediate intonational phrase, which enters a weak-strong metrical relation with other intermediate phrases in the sentence, and receives the primary accent.
- This proposal requires further testing, but may help explain why subjects might extrapose in presentational contexts.
- It also helps to explain why extraposition targets particularly heavy DPs: these are the DPs, in Ladd's own observation, which are most likely to constitute their own intermediate phrase. Extraposition may be a way to syntactically facilitate this.

## Conclusion

- I have argued, based on quantitative data from a parsed corpus of ENHG, that there are both similarities and differences between subject and object extraposition in ENHG.
  - Both subjects and objects may be extraposed to express narrow focus on the extraposed DP.
  - However, subjects may also be extraposed for another motivation: as a means to achieve default accent on the subject, particularly in presentational contexts.
  - As a result, subject extraposition occurs more frequently with quantified subjects, as well as with entities are new to the discourse.
- The result is therefore twofold:
  - 1 Subjects may have multiple motivations to extrapose (that is, either to obtain narrow focus or to obtain default accent in a presentational context).
  - 2 The phenomenon of DP extraposition in ENHG apparently demonstrates a subject-object asymmetry.

## Implications

- This has certain consequences for the broader study of information structure today.
  - Subject extraposition in ENHG shows that there is not necessarily a one-to-one correspondence between syntactic construction and information structural interpretation; rather, in this case, the syntax may be manipulated to accomplish multiple information structural goals (cf. Féry, 2007).
  - DP extraposition in ENHG also fits into a broader set of crosslinguistic focus phenomena which demonstrate a subject-object asymmetry (cf. Hartmann and Zimmermann, 2007; Skopeteas and Fanselow, 2010), raising important questions about the relationship between argument structure and information structural notions.
- On an unrelated note, the current study demonstrates the possibility of using parsed corpora as resources in information structural research, and on the importance of quantitative data when exploring such subtle and complex issues.

## Acknowledgements

Thanks to Anthony Kroch, Joel Wallenberg, Ivona Kucerova and Aaron Ecaj for their help and advice on this paper. Thanks also to the attenders of the 36th Annual Penn Linguistics Colloquium (PLC) for their feedback on an earlier version of this work.  
All errors are, naturally, my own.



## Bibliography I

- Bies, Ann (1996). “Syntax and Discourse Factors in Early New High German: Evidence for Verb-Final Word Order”. MA thesis. University of Pennsylvania.
- Bikel, Daniel M. (2004). “On the parameter space of lexicalized statistical parsing models”. PhD thesis. Department of Computer and Information Sciences, University of Pennsylvania.
- Diesing, Molly (1990). “Verb Movement and the Subject Position in Yiddish”. In: *Natural Language & Linguistic Theory* 8.1, pp. 41–79.
- Féry, Caroline (2007). “Information structural notions and the fallacy of invariant correlates”. In: *Interdisciplinary studies on information structure* 6.
- Frey, Werner (2006). “Contrast and movement to the German prefield”. In: *The architecture of focus (Studies in Generative Grammar 82)*. Ed. by V. Molnar and S. Winkler. Berlin, New York: Mouton de Gruyter, pp. 235–264.

## Bibliography II

- Hartmann, K. and M. Zimmermann (2007). “In Place - Out of Place? Focus in Hausa”. In: *On Information Structure: Meaning and Form*. Ed. by K. Schwabe and S. Winkler. Amsterdam: John Benjamins, pp. 365–403.
- Kroch, Anthony, Beatrice Santorini, and Ariel Diertani (2004). *Penn-Helsinki Parsed Corpus of Early Modern English*. <http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html>.
- Kroch, Anthony and Ann Taylor (2000). *Penn-Helsinki Parsed Corpus of Middle English, second edition*. <http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3/index.html>.
- Ladd, D. Robert (1996). *Intonational phonology*. Vol. 79. Cambridge studies in linguistics. United Kingdom: Cambridge University Press.
- Prince, Ellen (1989). “Yiddish Wh-clauses, subject-postposing, and topicalization”. In: *ESCOL '88*. Ed. by J. Powers and K. de Jong. Columbus, OH: Ohio State University, pp. 403–15.

## Bibliography III

- Skopeteas, S. and G. Fanselow (2010). “Focus types and argument asymmetries: a cross-linguistic study in language production”. In: *Contrastive information structure*. Ed. by C. Breul and E. Göbbel. Benjamins, pp. 169–197.
- Taylor, Ann et al. (2003). *The York-Toronto-Helsinki Parsed Corpus of Old English Prose*. University of York.
- Truckenbrodt, Hubert (2007). “The syntax-phonology interface”. In: *The Cambridge handbook of phonology*, pp. 435–456.
- Vallduví, Enric (1992). *The informational component*. New York: Garland.