The Loss of Bare Singular Noun Phrases in the History of English

Kenneth Hanson, Cristina Schmitt, and Alan Munn

Department of Linguistics and Languages
Michigan State University

April 11, 2014

Bare Noun Phrases

Mass nouns and plural count nouns in English do not require an article, while singular count nouns do.

(1) a.   Mary saw ∅ water on the floor.

  b.   Mary saw ∅ crumbs on the floor.

  c. * Mary saw ∅ penny on the floor.

Many pre-modifiers besides articles fill the need for something to precede a singular (count) noun.

(2)   Mary saw $\left\{ \begin{array}{c} a \\ one \\ the \\ this \\ every \\ John's \end{array} \right\}$ penny on the floor.

Bare Noun Phrases

Definition

Bare noun phrase: a noun phrase that is not modified by...

- an article (a/an, the)
- a numeral (one, two, ...)
- a demonstrative (this, that, these, those)
- a quantifier (every, some, many, ...)
- a genitive possessor (my, his, John's, ...)

Bare Singulars in Middle English

The range of singular noun phrases that could be bare in Middle English (c. 1150–1500) was greater than in Modern English.

(3)   a.   And to +te kyng of Frauns he was grete enmye...
           (CMCAPCHR,131.3063)

      b.   ...as keper of a town, +tat was famed to hym as he hadde
           wastyd his goodes... (CMWYCSER,256.559)

      c.   +tere he was i-made preost of Agelbertus;
           (CMPOLYCH,VI,187.1349)

      d.   He sheweth there that Craesus, the riche king of Lidia, whom
           Cyrus had taken prisoner... (ELYOT-E1-H,154.162)

Arguments and Predicates

The singular nouns that appear bare in Middle English are typically predicates. Singular arguments cannot usually be bare in Middle English (ME) or Modern English (ModE).

(4)   a.   John saw doctor.   (Arguments: * ModE, * ME)
      b.   John is doctor.   (Predicates: * ModE, ✓ ME)

Definitions

argument noun phrase: a noun phrase that introduces a referent
- Includes subjects and objects of most verbs

predicate noun phrase: a noun phrase that denotes a property
- Includes be-predicates, as-predicates, small clause predicates, secondary predicates

Role Predicates

Not all singular predicates can be bare in Middle English. Those that are bare are typically role predicates.

    (5)   This is chair.   (Non-roles: * ModE, * ME)

    (6)   John is doctor.   (Roles: * ModE, ✓ ME)

Unique roles predicates are exceptionally bare among singular nouns in Modern English.

    (7)   John is president.   (Unique roles: ✓ ModE, ✓ ME)

Role Predicates: Definitions

role noun a noun that denotes a role in some domain, as opposed to an inherent property of an individual

- Professions: doctor, lawyer, teacher, plumber, ...
- Clergy: monk, nun, pastor, ...
- Other roles: master, servant, keeper, ...

unique role noun a noun that denotes a role stereotypically filled by a single individual in some domain

- Political offices: president, governor, king, duke, ...
- High ranking clergy: pope, bishop, archbishop, ...

role predicate a role noun (phrase) in a predicative position

- ex. John is a doctor.

unique role predicate a unique role noun (phrase) in a predicative position

- ex. John is president.

The History of English

- c. 500–1150: Old English (OE)
    - c. 750–1050: Composition of Beowulf
    - 1066: Norman Conquest
- 1150–1500: Middle English (ME)
    - 1343–1400: Chaucer
- 1500–1750: Early Modern English (EME)
    - 1564–1616: Shakespeare
- 1750–Present: Modern English (ModE)
    - 1755: Publication of Samuel Johnson's *A Dictionary of the English Language*

Summary and Questions

Three basic stages of English:

OE No articles Most nouns bare	→	**ME/EME** Def+Indef articles Some BSNPs	→	**ModE** Only unique roles can be bare

Two questions:

1. What is the distribution of bare singular noun phrases in ME?
2. What changed between ME and ModE?
   - Note: For changes between OE and ME, cannot easily disentangle loss of bare nouns phrases from the rise of the article system.

Basis of the Study

Our hypothesis: English underwent (at least one) parametric shift in its grammar that resulting in the lost of most bare singular noun phrases.

- We can test this using the Constant Rate Effect (Kroch 1989).

Constant Rate Effect

For grammatical changes, the rate of change in frequency of use over time will be constant across contexts.

Our approach: a corpus study.

- Data will be automatically collected from a series of parsed corpora.
- The data will be coded for syntactic position and lexico-semantic class so that we can track the distribution of bare singulars over time in different contexts.

Data Sources: English Historical Corpora

- Penn-Helsinki Parsed Corpus of Middle English, 2nd Edition (PPCME2)
  - Time period: c. 1150-1500
- Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)
  - Time period: c. 1500-1710
- York-Helsinki Parsed Corpus of Early English Correspondence (PCEEC)
  - Time period: c. 1410-1695

Ex. One PPCME2 Token

```
( (IP-MAT (NP-SBJ (D +te) (N blowere)
                  (IP-INF (TO to)
                          (VB melte)
                          (NP-OB1 (D +tys) (N gold))))
        (BEP is)
        (NP-OB1 (D +te) (NPR deuel))
        (. ,))
   (ID CMAELR3,27.31))
```

Overview

Basic steps:

1. Create a list of singular count nouns to track.
2. Search the corpora for all noun phrases with head nouns in this list.
3. Code each noun phrase for semantic class, syntactic position, date of source text, and other relevant properties.
4. Analyze the results using the R statistical programming language.

Items to Track

Class	Lexical Items
Unique Roles	bishop, duchess, duke, emperor, empress, king, lord, pope, queen
Non-Unique Roles	ally, doctor, friar, guard, keeper, knight, monk, nun, priest, prince, princess, prisoner, servant, sinner, slave, soldier, teacher
Non-roles (Animate)	beast, ghost, maiden, man, woman
Non-roles (Concrete)	body, book, face, house, mouth, voice, weapon
Non-roles (Abstract)	deed, end, heart, name, soul, tale, thing, word, world

- Lexical items were chosen on the basis of appearing throughout the time period under investigation, and not having word senses in more than one of our classes.
- A list of all spelling variants for each noun was compiled.
- Non-roles were collapsed for analysis.

Searching and Coding with CorpusSearch
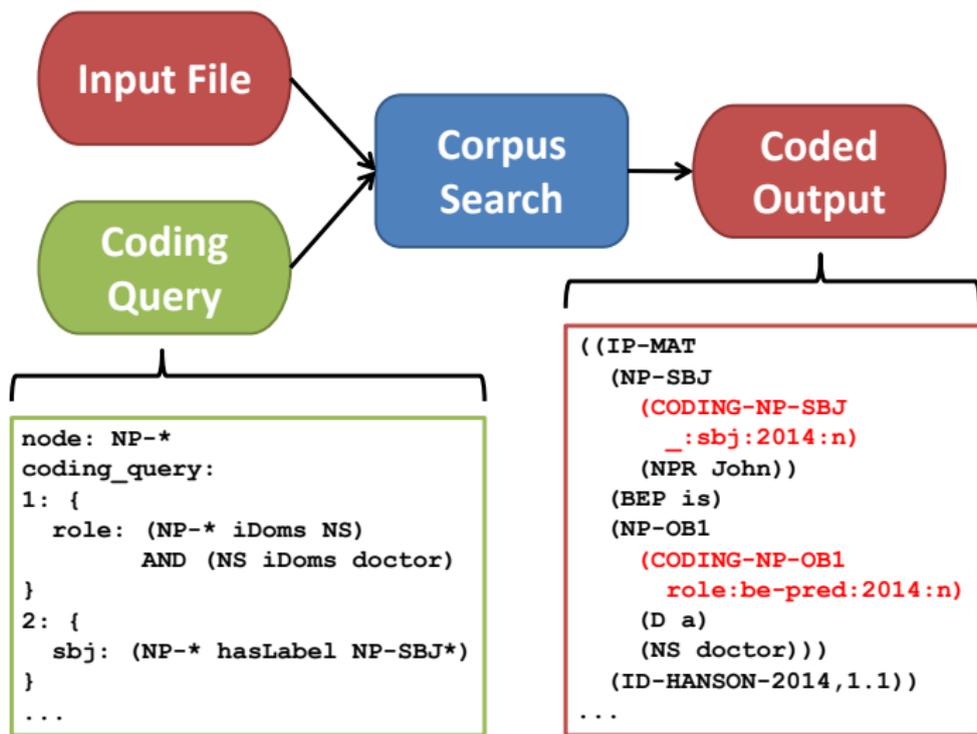
Multi-step Search and Coding Process
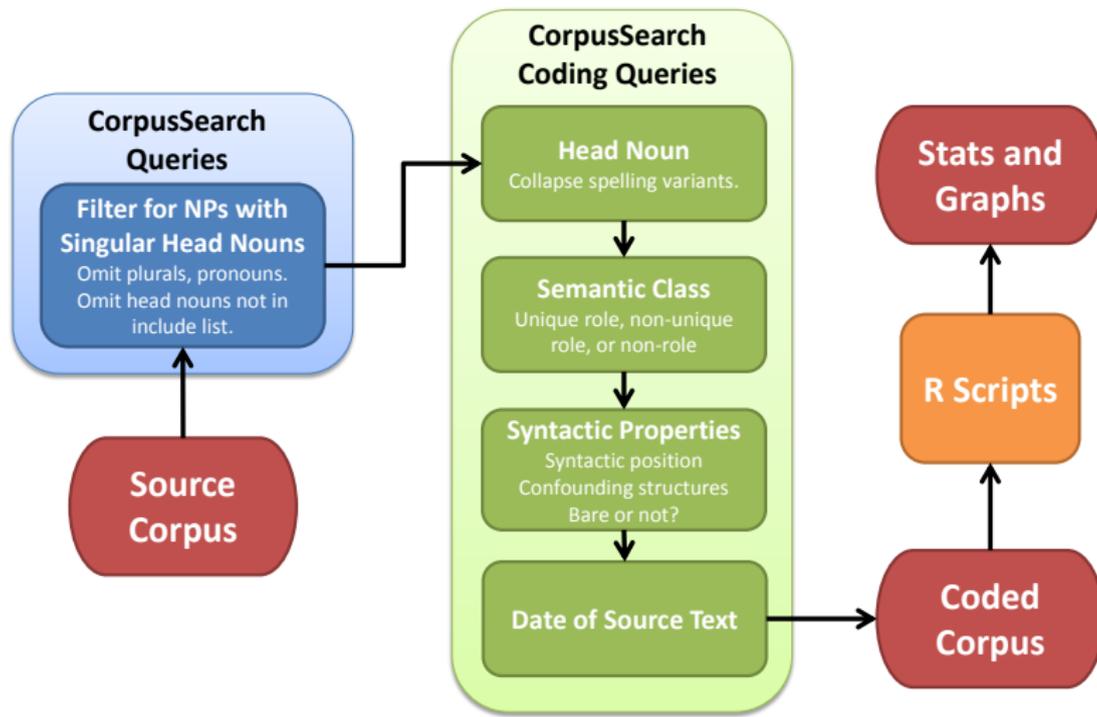
Post-Processing and Analysis with R

Data cleanup:

- Filtered out confounding/irrelevant data tokens.
- Collapsed coding values to remove irrelevant contrasts in coding scheme.

Statistical analysis:

- Ran (GLM) logistic regressions of the frequency of bareness over time, partitioned by semantic class and syntactic position.

Data Token Counts

Period	Class	As-Pred.	Be-Pred.	SC/SP	Verb Arg.
ME	non-role	44	226	28	5812
ME	role-non-uniq	8	75	42	323
ME	role-uniq	12	448	185	1722
EME	non-role	89	298	21	5688
EME	role-non-uniq	10	109	48	472
EME	role-uniq	4	42	27	1085

Some caveats:

- Some regression curves will have large confidence intervals.
- When the center portion of the S-curve is not within our time period, we must be cautious when drawing conclusions about the model.
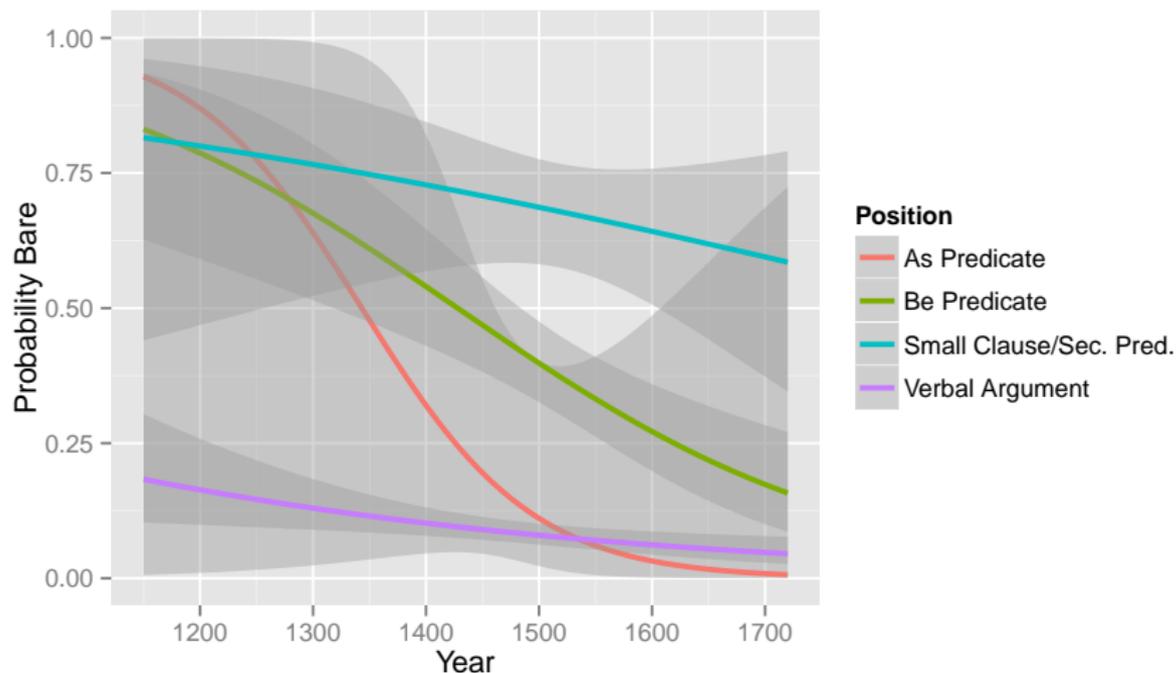
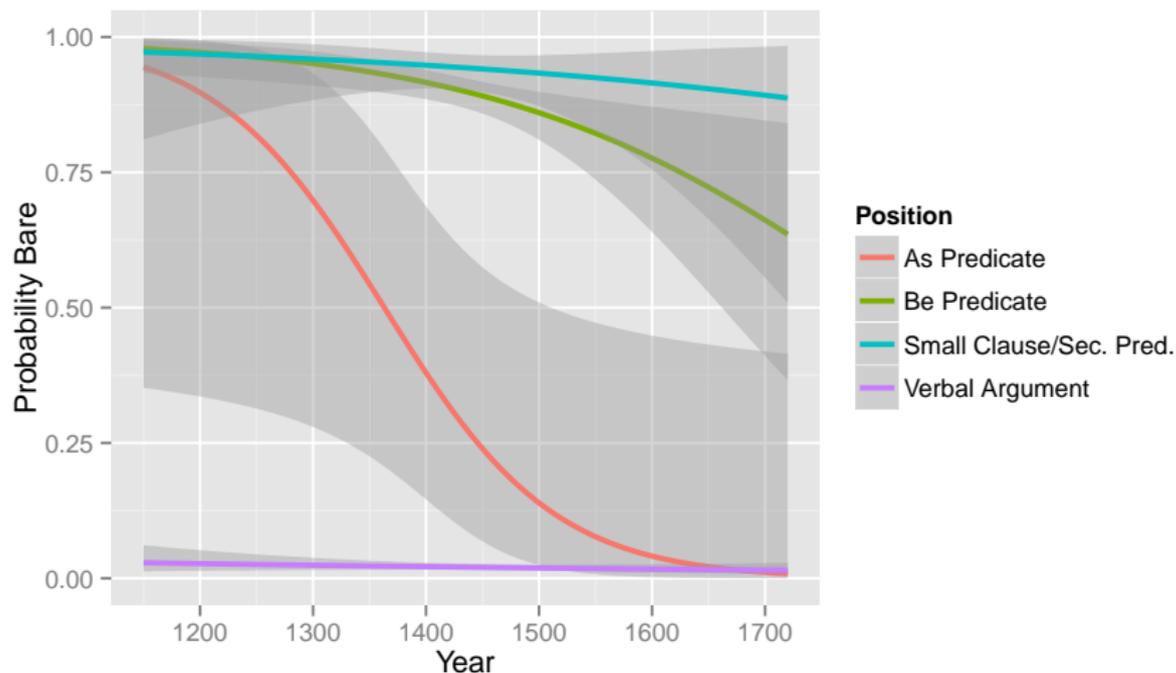# All Nouns by Semantic Class
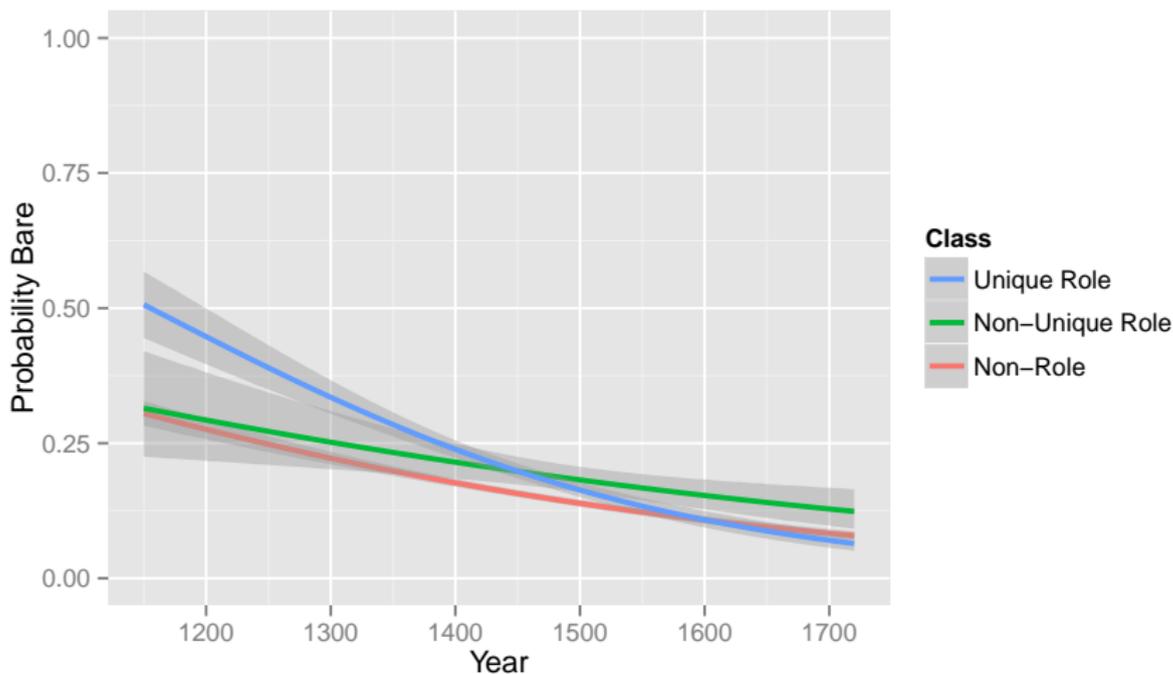
All Nouns by Syntactic Position
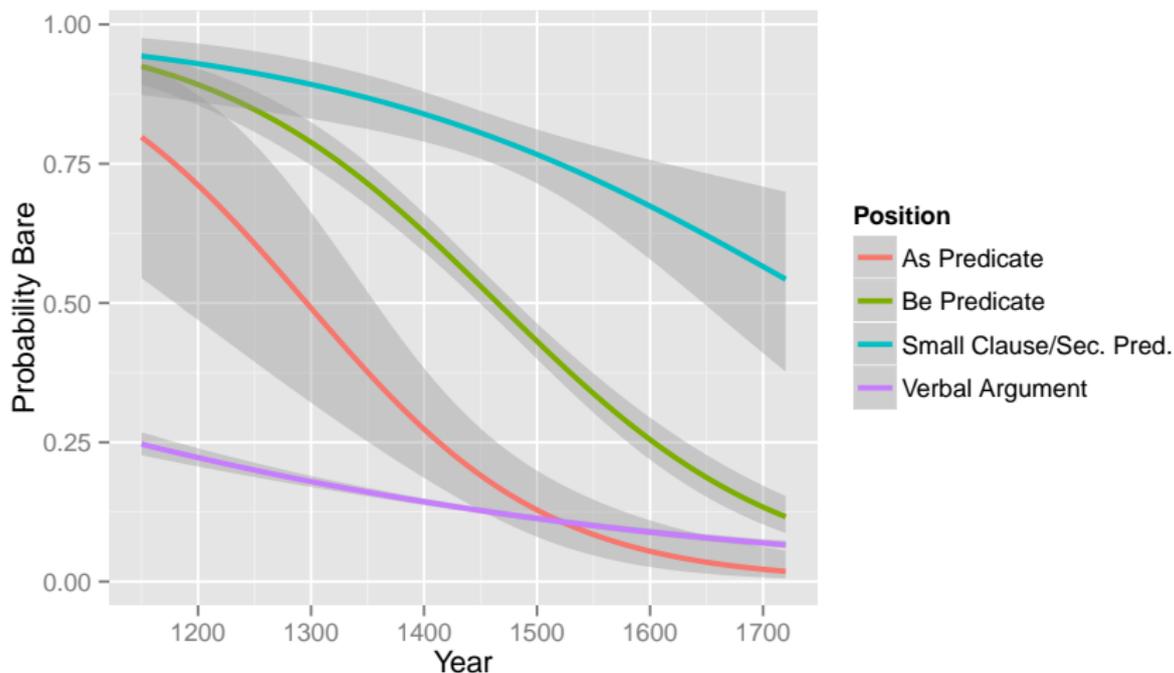
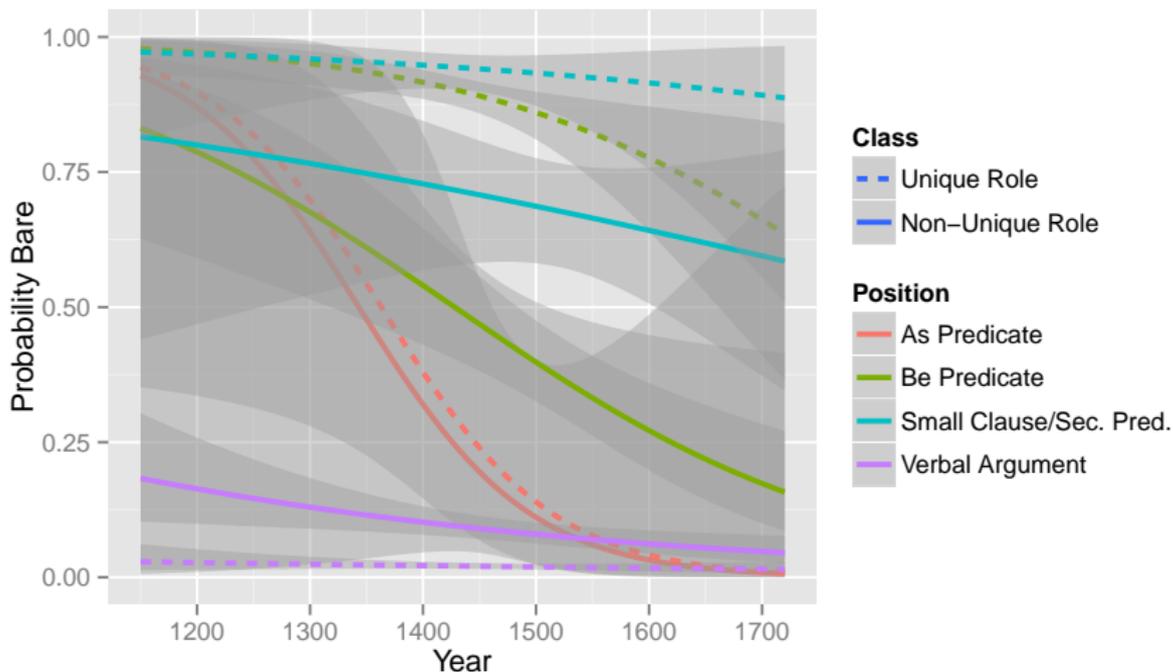# Non-Roles by Position
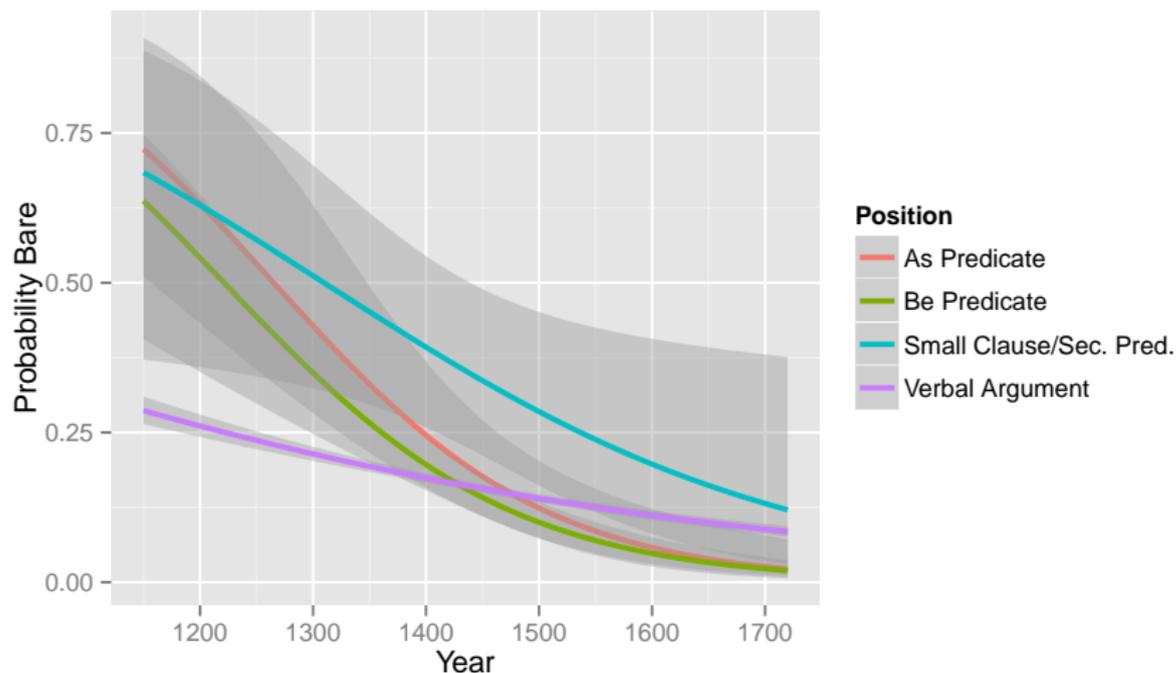
Non-Unique Roles by Position

Unique and Roles by Position

Unique and Non-Unique Roles by Position

Summary

- Overall trends of bare frequency:
  - Unique roles > Non-unique roles > Non-roles
  - SC/SP > Be-predicates > As-predicates > Arguments
  - All categories decreased over time.
- Bare arguments of all classes were already infrequent at the beginning of the Middle English period.
- Bare predicates were frequent in early Middle English. Non-roles were already declining, and non-unique roles began to decline soon after.
  - As-predicates appear to behave separately from the other groups, but the number of data points is very small.
- Unique role predicates declined somewhat over the period of investigation, but remain possible in Modern English.

Acknowledgements

References I

1. Alan Munn and Cristina Schmitt (2005). Number and indefinites. Lingua, 115(6), 821-855.

2. Kroch, Anthony (1989). Reflexes of grammar in patterns of language change. Language variation and change, 1(03), 199-244.

3. Anthony Kroch and Ann Taylor (2000). The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, (http://www.ling.upenn.edu/hist-corpora/).

4. Anthony Kroch, Beatrice Santorini, and Lauren Delfs (2004). The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME). Department of Linguistics, University of Pennsylvania. CD-ROM, first edition, (http://www.ling.upenn.edu/hist-corpora/).

References II

5 Ann Taylor, Arja Nurmi, Anthony Warner, Susan Pintzuk, and
  Terttu Nevalainen (2006). The York-Helsinki Parsed Corpus of Early
  English Correspondence (PCEEC). Department of Linguistics,
  University of York. Oxford Text Archive, first edition,
  (http://www-users.york.ac.uk/ lang22/PCEEC-manual/index.htm).

6 Beth Randall (2010). CorpusSearch
  (http://corpussearch.sourceforge.net/index.html).