Modeling interactions in morphosyntactic changes

Hezekiah Akiva Bacovcin

University of Pennsylvania - Department of Linguistics

Introduction

GOAL Give an example of two changes in a bleeding relationship, and demonstrate a model for their interaction.

- **Bleeding** - where a later change removes the context where a previous (or ongoing) change applies

Data

- Penn Parsed Corpora of Historical English (PPCME2, PPCEME, PPCMBE)
- Parsed Corpus of Early English Correspondence (PCEEC)
- York Corpus of Old English Prose (YCOE)

Old English - Active

IO–DO Ic sealde him þine spræce.
I.NOM give.PST 3PL.**DAT** 2SG.POSS speech.ACC
'I gave them your speech.' (John 17:14 West Saxon Gospel)

DO–IO ... he sealde hyne Iudas Scariothe.
... he.NOM give.PST it.ACC Judas.**DAT** Iscariot.**DAT**
'... he gave it to Judas Iscariot.' (John 13:26 West Saxon Gospel)

Modern English - Active

IO–DO
- I gave John the books
- * I gave **to** John the books
- I gave **to** John the book that I wrote last summer.

DO–IO
- * I gave the books John
- I gave the books **to** John

McFadden (2002)

- Only looked at the PPCME2

McFadden (2002)

- Only looked at the PPCME2
- Proposed two different structures in Old English for different word orders
- The change of dative case to 'to' occurs only in the DO–IO word order

Problem

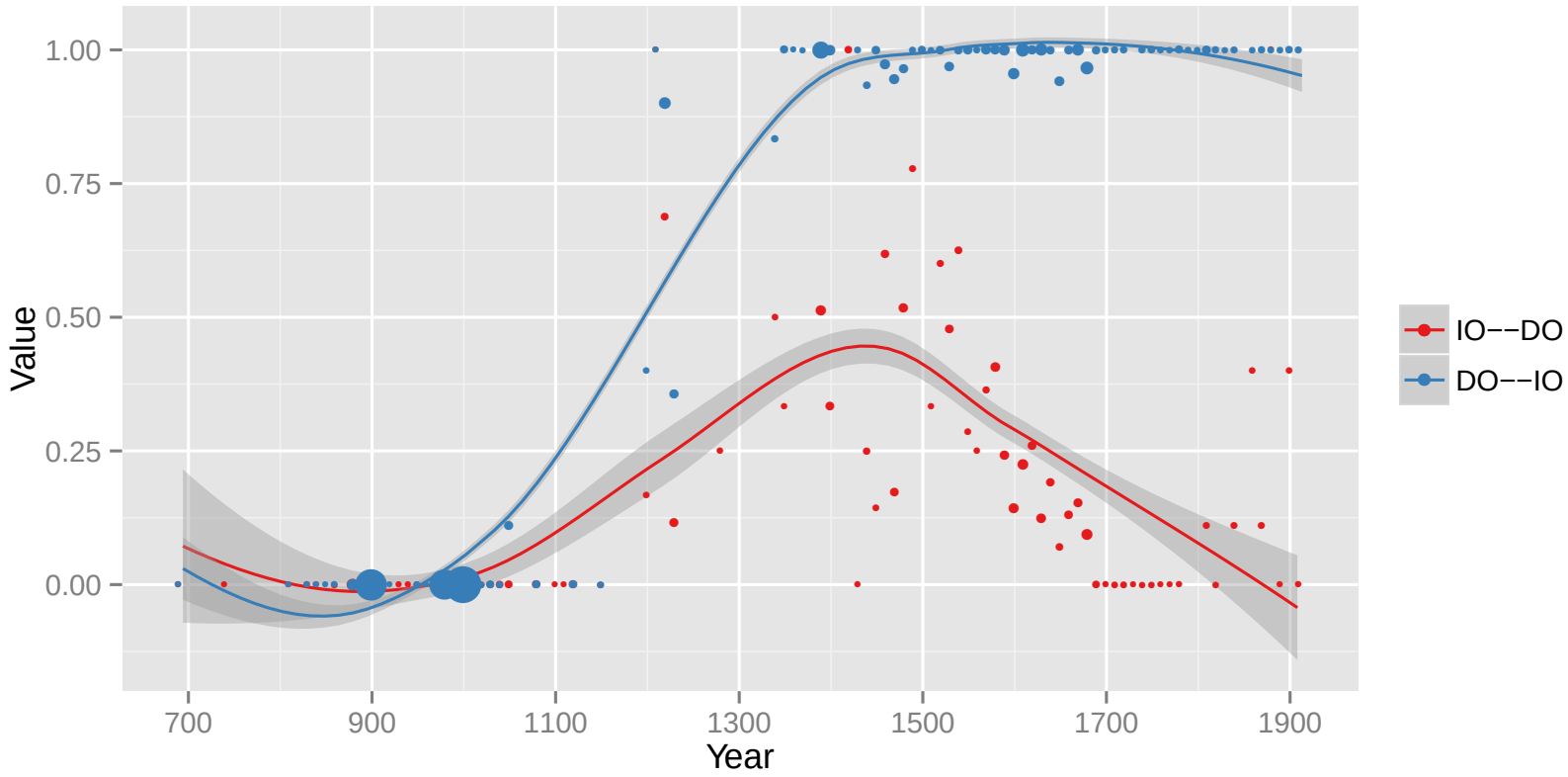

Figure : Percent of sentences with a non-pronominal Indirect Object marked with 'to' (using a LOESS regression)

Solutions

- McFadden suggests that this rise can be accounted for by differences in **heavy NP shift**, due to changes in text complexity.

Solutions

- McFadden suggests that this rise can be accounted for by differences in **heavy NP shift**, due to changes in text complexity.
- While the later texts do have longer noun phrases than the earlier ones, this does not account for the rise of 'to' in the IO–DO word order.
- I suggest that the decline in the IO–DO order is do to bleeding from a different change that happened around 1400.

Passive Possibilites (Allen, 1999)

- Dative fronted passives (as seen Old English) end around the loss of dative case.
- No clear examples of recipient passives (i.e. with a pronominal recipient, where case differences can be identified) can be found until 1375.

Old English - Passive

IO–DO him wearð geseald an snæd
him.**DAT** be.PST.3SG given one.ACC slice.ACC
flæsces.
meat.GEN
'He was given a slice of meat.' (Ælfric's Homilies)

DO–IO he byð þeodum geseald.
he.NOM be.3SG nation.**DAT**.PL given
'He will be delivered to the nations.' (Luke 18:32 West Saxon Gospel)

Modern English - Active

IO–DO John was given the books.

DO–IO The books were given *(to) John.

Corpus Findings

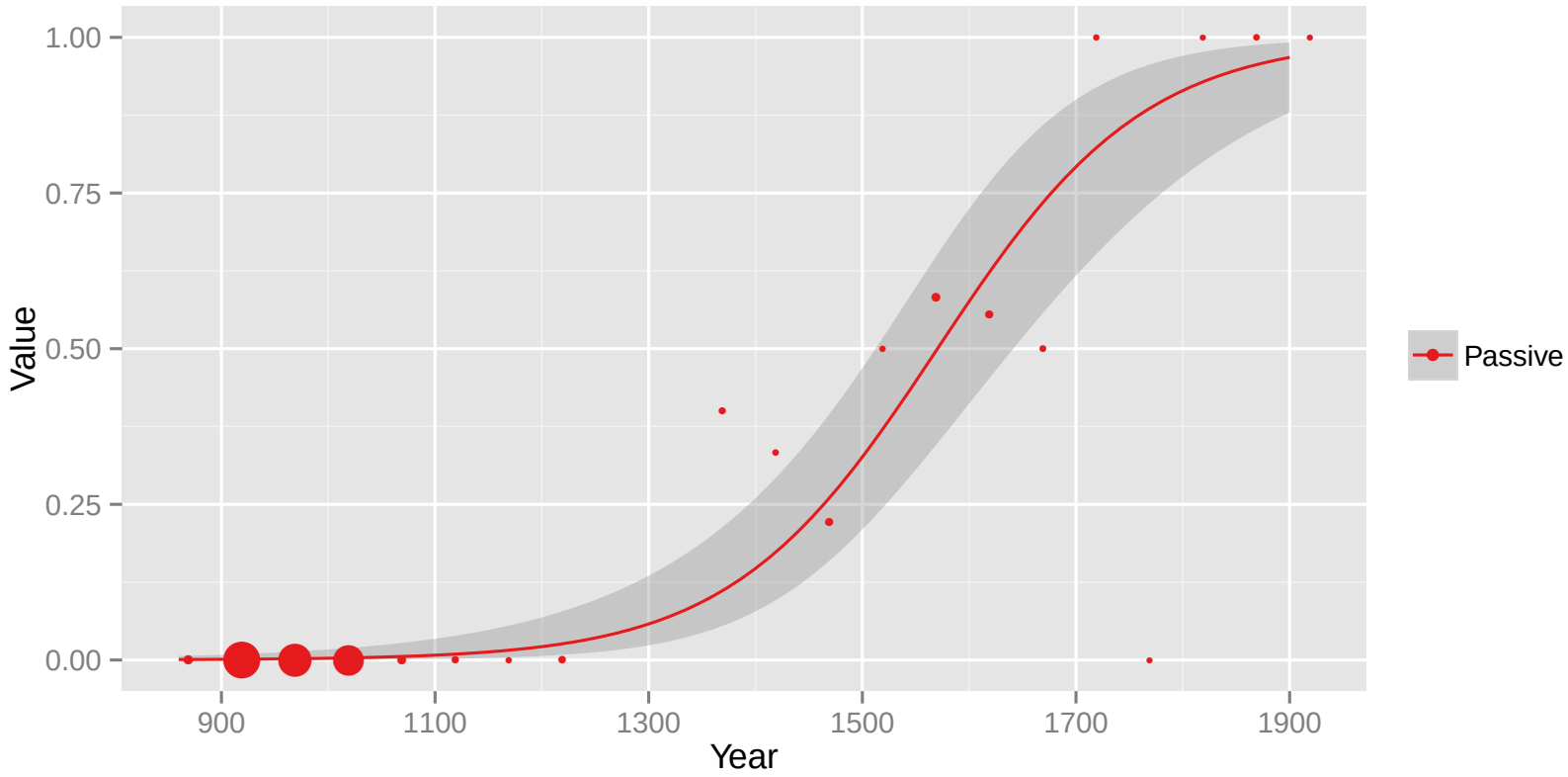

Figure : Percent of sentences, without 'to' marking, that have non-pronominal recipient subjects (Using a logistic regression)

Analysis

- Recipient passivisation becomes possible when the recipient is reanalysed as receiving structural case.
- At some point in the 14th century there is a reanalysis of the IO–DO word order from being a **dative–accusative** construction to an **accusative–accusative** construction
- 'to' enters in all word orders as the new dative case marker
- The loss of dative assignment in the IO–DO word order **bleeds** the application of 'to' dative marking.

Examples

- I.NOM gave John.**ACC** the books.ACC.
- John.**NOM** was given the books.ACC.
- I.NOM gave the books.ACC **to** John (DAT).
- The books.NOM were given **to** John (DAT).

Logistic

- The general statistical model for linguistic change has been the logistic, which models the spread of a feature through a population
- The basic form of the logistic is $\frac{1}{1+e^{a+bx}}$
- **a** is the intercept (i.e. where the 50% mark is)
- **b** is the slope (i.e. how fast the the change moves away from the 50% mark)

Constant Rate Hypothesis (Kroch, 1989)

- Hypothesis: A single change propagating through different grammatical environments will have the same slope in each environment, but may have different intercepts.

Logistic Model for Bleeding

- Bleeding - Rule 1 can only apply when rule 2 has not yet bled its environment

Logistic Model for Bleeding

- Bleeding - Rule 1 can only apply when rule 2 has not yet bled its environment
- $p_{rule1} = \frac{1}{1+e^{a_1+b_1 x}}$
- $p_{rule2} = \frac{1}{1+e^{a_2+b_2 x}}$

Logistic Model for Bleeding

- Bleeding - Rule 1 can only apply when rule 2 has not yet bled its environment
- $p_{rule1} = \frac{1}{1+e^{a_1+b_1 x}}$
- $p_{rule2} = \frac{1}{1+e^{a_2+b_2 x}}$
- $p_{bledrule}$ =Rule 1 * (1 - Rule 2) [i.e. the chance of having rule 1 apply * the chance of having rule 2 not apply]

DO–IO Equation

- $p_{\text{DO–IO}} = \frac{1}{1+e^{a_{\text{DO–IO}}+b_{\text{DO–IO}}x}}$
- Parameters:
  1. $a_{\text{DO–IO}}$ = Intercept for DO–IO 'to' dative
  2. $b_{\text{DO–IO}}$ = Slope for DO–IO 'to' dative

Passive Equation

- $p_{pas} = \frac{1}{1+e^{a_{pas}+b_{pas}x}}$
- Parameters:
  1. a_{pas} = Intercept for recipient passivisation
  2. b_{pas} = Slope for recipient passivisation

IO–DO Equation, Part I

- $p_{\text{IO–DO (to)}} = \frac{1}{1+e^{a_{\text{IO–DO (to)}}+b_{\text{IO–DO (to)}}x}}$
- Parameters:
  1. $a_{\text{IO–DO (to)}}$ = Intercept for IO–DO 'to' dative
  2. $b_{\text{IO–DO (to)}}$ = Slope for IO–DO 'to' dative

IO–DO Equation, Part II

- $p_{\text{IO–DO (acc)}} = \frac{\textbf{heavy}}{1+e^{a_{\text{IO–DO (acc)}}+b_{\text{IO–DO (acc)}}x}}$
- Parameters:
  1. heavy = Rate of heavy NP shift (creating 'to' marked IO–DO sentences from DO–IO bases)
  2. $a_{\text{IO–DO (acc)}}$ = Intercept for IO–DO accusative reanalysis
  3. $b_{\text{IO–DO (acc)}}$ = Intercept for IO–DO accusative reanalysis

IO–DO Equation, Part III

- $p_{\text{IO–DO}} = p_{\text{IO–DO (to)}} * (1 - p_{\text{IO–DO (acc)}})$

Hypothesis

- Assuming the change in the IO–DO word order is just the interaction of the 'to' dative change and the accusative reanalysis change, the Constant Rate Hypothesis predicts:
  1. $b_{\text{IO–DO (to)}} = b_{\text{DO–IO}}$
  2. $b_{\text{IO–DO (acc)}} = b_{\text{pas}}$
  3. All of the intercepts ($a_{\text{DO–IO}}$, etc.) can be different

How to Estimate the Parameters

- For DO–IO and passive (i.e. equations without multiplication), standard logistic regression techniques can be used to estimate the values.
- For the IO–DO, the interaction between IO–DO (to) and IO–DO (acc) prohibit the use of standard logistic regression techniques.

How to Estimate the Parameters

- For DO–IO and passive (i.e. equations without multiplication), standard logistic regression techniques can be used to estimate the values.
- For the IO–DO, the interaction between IO–DO (to) and IO–DO (acc) prohibit the use of standard logistic regression techniques.
- Instead, a Markov chain Monte Carlo (MCMC) simulation can be used to estimate these values.

Markov chain Monte Carlo

- A random walk through the parameter space, where the probability of changing the value of a particular parameter at a given **iteration** is related to probability that it is the correct modeling the data.
- Since walk is more likely to linger on the correct values, the distribution of the iterations should best estimate the probability distribution of the parameters given the data.
- Useful for estimating values that do not have analytic solutions.

MCMC with RStan (Stan Development Team, 2013)

- Multiple chains are run to avoid finding local maxima.
- Each iteration has different values for all parameters.
- A number of iterations are run and not saved in order to eliminate influence from the initialisation settings (**Burn-in**)
- After burn-in, not every iteration is saved in order to avoid clumping.

Examining the Output

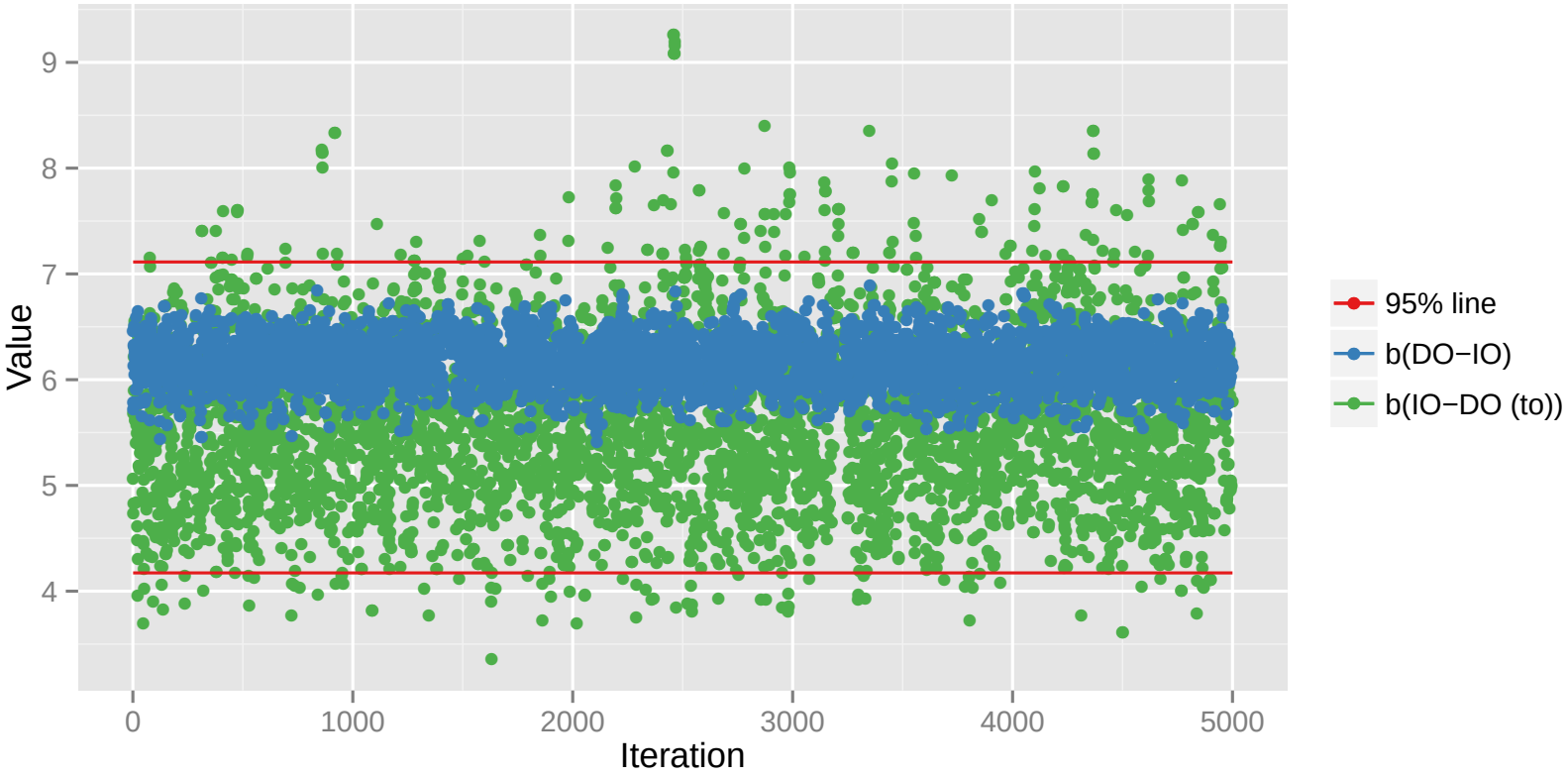

Figure : Values for the b_{DO-IO} and $b_{IO-DO\ (to)}$ parameter in each iteration

Difference between $b_{DO–IO}$ and $b_{IO–DO\ (to)}$

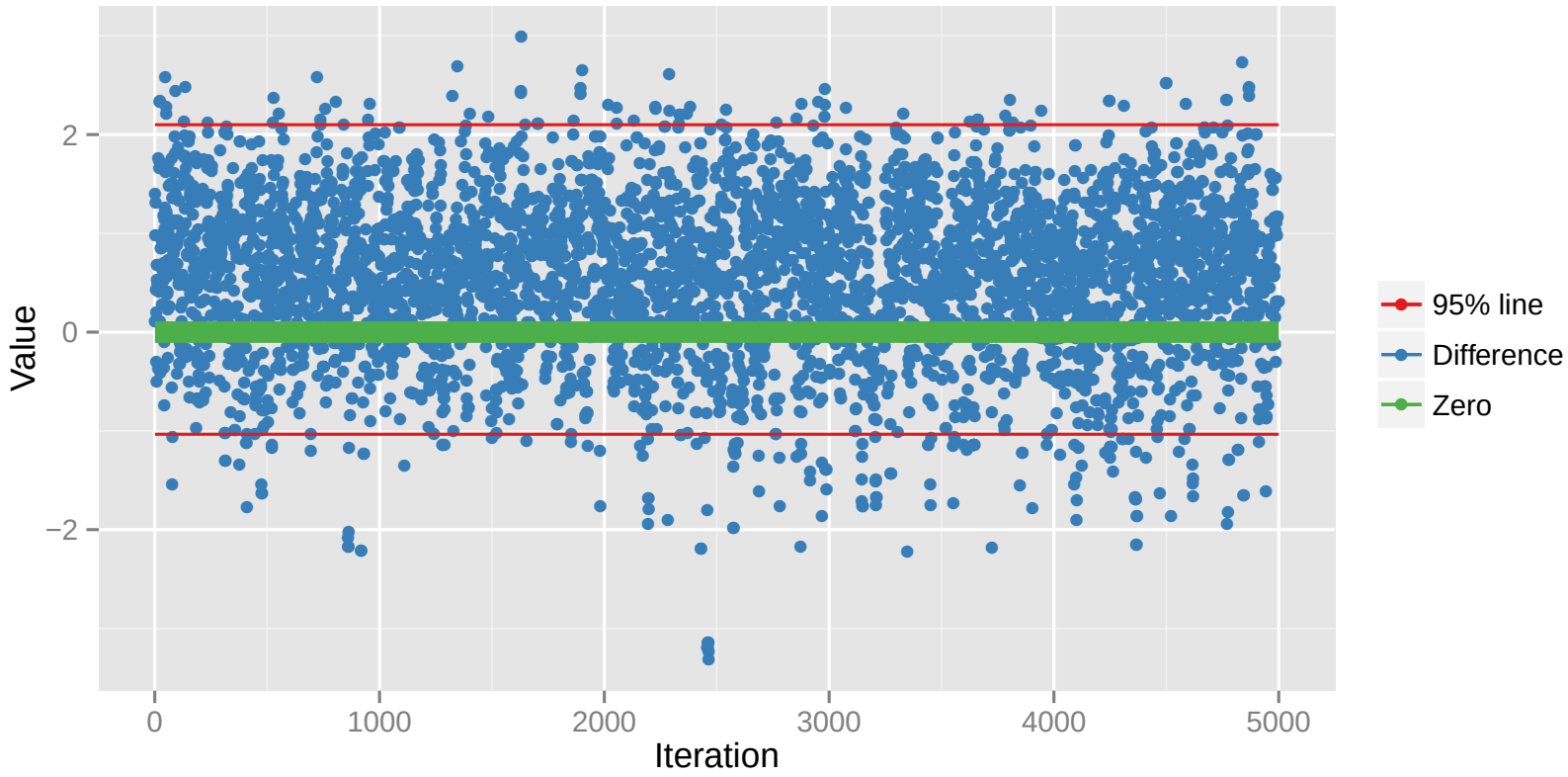

Figure : Value of $b_{DO–IO}$ - $b_{IO–DO\ (to)}$ in each iteration

Results Difference between b_{pas} and $b_{IO-DO(acc)}$

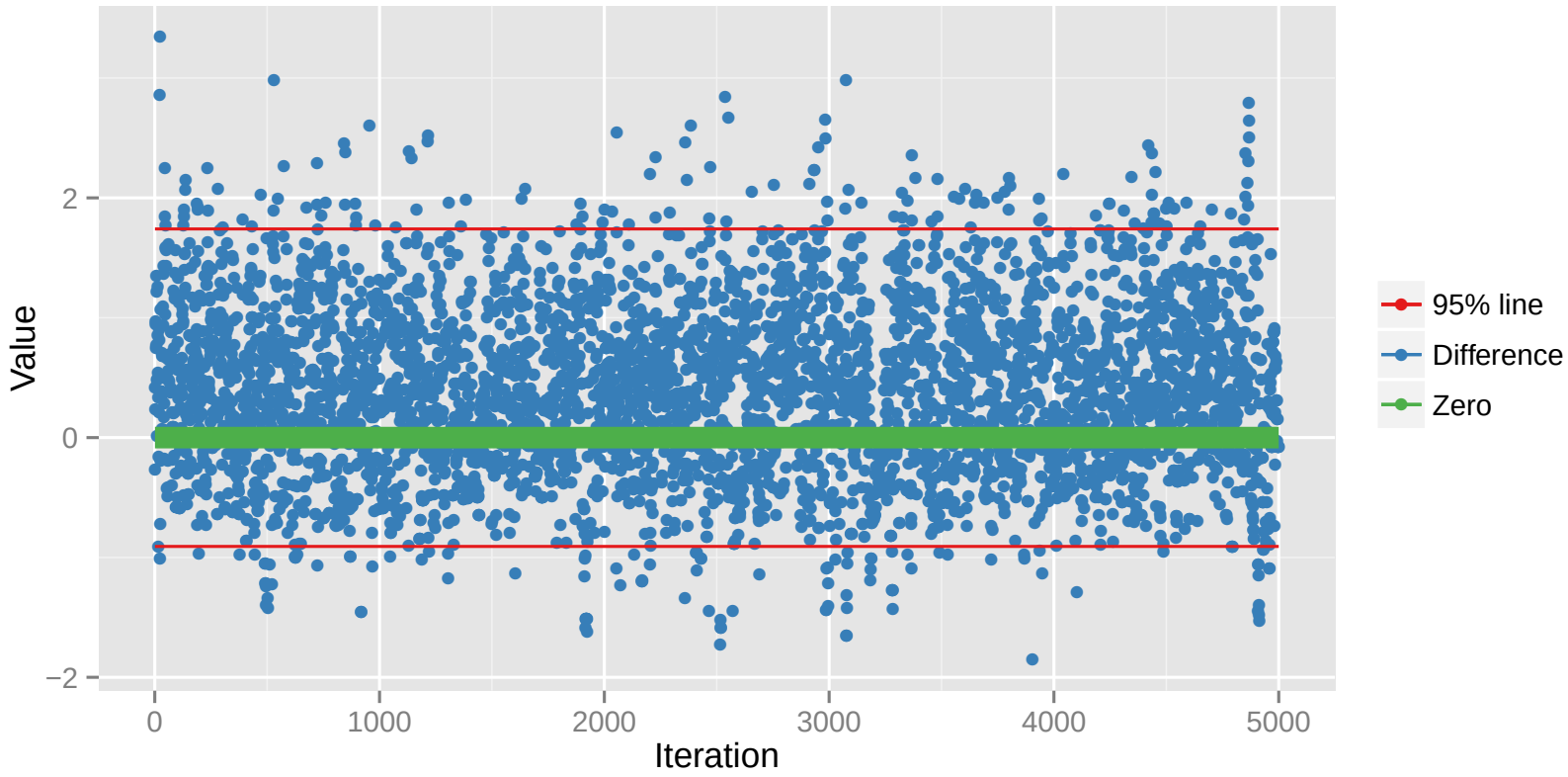

Figure : Value of $b_{IO-DO(acc)}$ - b_{pas} in each iteration

Comparing MCMC and LOESS

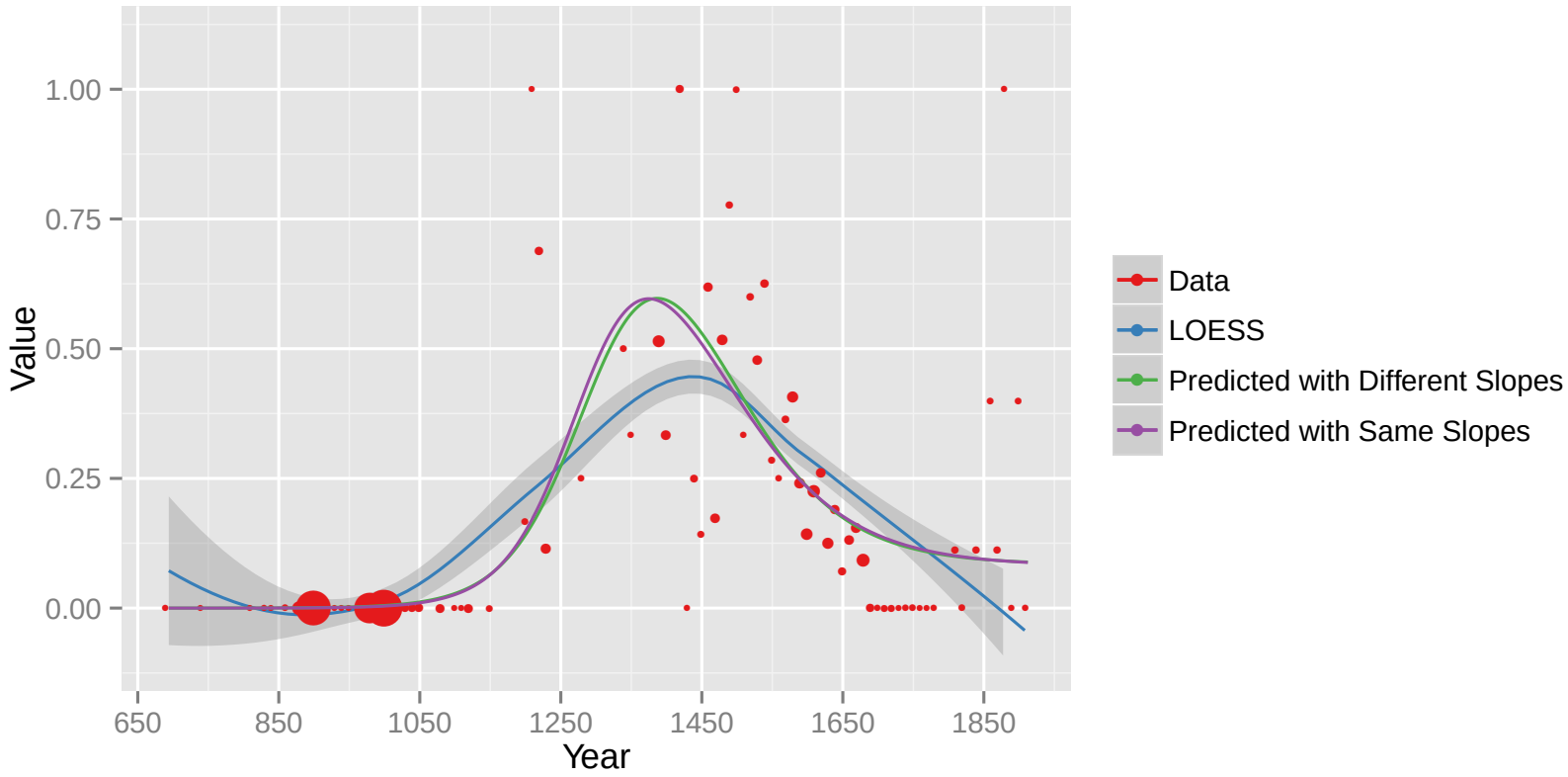

Figure : Comparison of Logistic Values from MCMC (Two Ways) and the LOESS prediction

Conclusions

- The data is consistant with the rate of ‘to’ dative in the IO-DO order coming from an interaction between a general rise in the use of the ‘to’ dative and accusative reanalysis
- Multiplication of independent logistics can produce the rise-fall pattern seen in a possible case of bleeding
- Markov chain Monte Carlo simulations can be useful for estimating the parameter values of equations lacking analytic solutions.

Further Work

- Gather more data, a more exact value of the parameters could be derived with more data.
- Finding other cases of interaction, which can be identified by the rise-fall tradjectory of the change.

Thank you to Tony Kroch and Beatrice Santorini and my fellow graduate students at the University of Pennsylvania for their help with this project.

Thank you very much for listening.

The scripts and queries used in this presentation can be found on my website at www.ling.upenn.edu/∼bacovcin.

References I

Allen, Cynthia L. 1999. *Case marking and reanalysis: Grammatical relations from old to early modern english*. Oxford University Press.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language variation and change* 1:199–244.

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

Kroch, Anthony, Beatrice Santorini, and Ariel Diertani. 2010. The Penn-Helsinki Parsed Corpus of Modern British English (PPCMBE). CD-ROM, first edition. URL `http://www.ling.upenn.edu/hist-corpora/`.

References II

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

Kruschke, John. 2010. *Doing bayesian data analysis: A tutorial introduction with r*. Academic Press.

McFadden, Thomas. 2002. The rise of the to-dative in middle english. *Syntactic effects of morphological change. Oxford* .

R Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`.

Stan Development Team. 2013. Stan: A c++ library for probability and sampling, version 1.3. URL `http://mc-stan.org/`.

References III

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

Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE). Oxford Text Archive, first edition. URL `http://www-users.york.ac.uk/~lang22/YcoeHome1.htm`.

Wickham, Hadley. 2009. *ggplot2: elegant graphics for data analysis*. Springer New York. URL `http://had.co.nz/ggplot2/book`.

Appendix: Heavy NP Shift

	AIC	BIC
Pronoun	99	113
Pronoun and CP Status	99	117
Pronoun and Length	100	118
Pronoun, Length and CP Status	100	124

Table : AIC and BIC comparison of IO-DO data before 1375 w.r.t DO weight

- "...and sente to kyng Cacanus worde" (John of Trevisa's Polychronicon, 1387)
- "in the same day thou shalt paye to a nedy traueylour his huyre" (Purvey's General Prologue to the Bible, 1388)