

# Applying Yang's model of language change: Resetting the headedness of IP in early English

Richard Zimmermann

Université de Genève

September 18, 2013

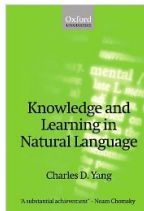
# Outline

Yang's theory of language change

Development of early English IP headedness

Applying Yang's theory to the change in IP headedness

Outlook



## Yang's theory of language change

Development of early English IP headedness

Applying Yang's theory to the change in IP headedness

Outlook

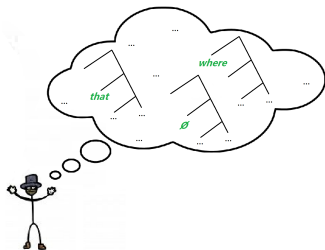
Variation

Replicator dynamics

How to calculate fitness

# Yang's theory of language change

# Variation as representation of multiple grammatical options



- (1)
- Where did he say **that** he went?
  - Where did he say  $\emptyset$  he went?
  - Where did he say **where** he went?
  - ...

## Terminology

What is the name for the source of (syntactic) variation?

## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)

## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)

# Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)



## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)

## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)
- ▶ (Holistic) constructions (CxG)

## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)
- ▶ (Holistic) constructions (CxG)
- ▶ (Holistic) variants (Sociolinguistics)

## Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)
- ▶ (Holistic) constructions (CxG)
- ▶ (Holistic) variants (Sociolinguistics)
- ▶ "Lingueme" (Croft 2000)

# Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)
- ▶ (Holistic) constructions (CxG)
- ▶ (Holistic) variants (Sociolinguistics)
- ▶ "Lingueme" (Croft 2000)
- ▶ **Yang: Competing grammar / Grammar** (Kroch 1989)

# Terminology

What is the name for the source of (syntactic) variation?

- ▶ Feature values on lexical items (Minimalist Program)
- ▶ Parameter settings (GB)
- ▶ OT-constraints (Optimality Theory)
- ▶ Annotated phrase structure rules (LFG)
- ▶ (Holistic) constructions (CxG)
- ▶ (Holistic) variants (Sociolinguistics)
- ▶ "Lingueme" (Croft 2000)
- ▶ **Yang: Competing grammar / Grammar** (Kroch 1989)
  - ⇒ the smallest (syntactic) unit of replication

## Yang's theory of language change

Development of early English IP headedness

Applying Yang's theory to the change in IP headedness

Outlook

Variation

**Replicator dynamics**

How to calculate fitness

# Relative frequency and fitness of a Grammar

## Relative frequency and fitness of a Grammar

- ▶ Adult speaker with two competing Grammars,  $G_1$  and  $G_2$



## Relative frequency and fitness of a Grammar

- ▶ Adult speaker with two competing Grammars,  $G_1$  and  $G_2$
- ▶ Each Grammar has a relative frequency:  
 $G_1$  has a proportion of  $p\%$ ;  $G_2$  has a proportion of  $q\%$

## Relative frequency and fitness of a Grammar

- ▶ Adult speaker with two competing Grammars,  $G_1$  and  $G_2$
- ▶ Each Grammar has a relative frequency:  
 $G_1$  has a proportion of  $p\%$ ;  $G_2$  has a proportion of  $q\%$
- ▶ All grammars are not equally "good" or "fit":  
 $G_1$  has an advantage of  $\alpha\%$ ;  $G_2$  has an advantage of  $\beta\%$

## Relative frequency and fitness of a Grammar

- ▶ Adult speaker with two competing Grammars,  $G_1$  and  $G_2$
- ▶ Each Grammar has a relative frequency:  
 $G_1$  has a proportion of  $p\%$ ;  $G_2$  has a proportion of  $q\%$
- ▶ All grammars are not equally "good" or "fit":  
 $G_1$  has an advantage of  $\alpha\%$ ;  $G_2$  has an advantage of  $\beta\%$
- ▶ The proportions of the Grammars in the following generation can be calculated using (standard) replicator equations:

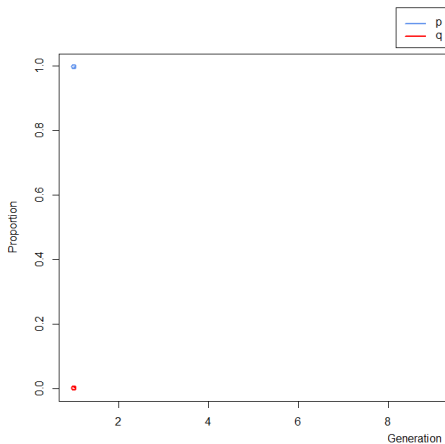
$$p_{t+1} = \frac{\alpha p_t}{\alpha p_t + \beta q_t}$$

# Dynamics of a two Grammar system

$$p = 1 \quad q = 0$$

$$\alpha = 0.1 \quad (\beta = 0.5)$$

$$p_1 = \alpha p / (\alpha p + \beta q)$$

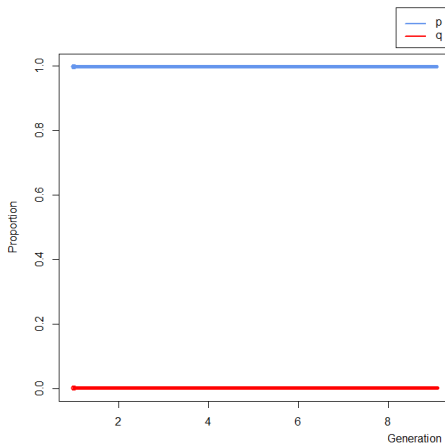


# Dynamics of a two Grammar system

$$p = 1 \quad q = 0$$

$$\alpha = 0.1 \quad (\beta = 0.5)$$

$$p_1 = 1 / (1 + 0) = 1$$

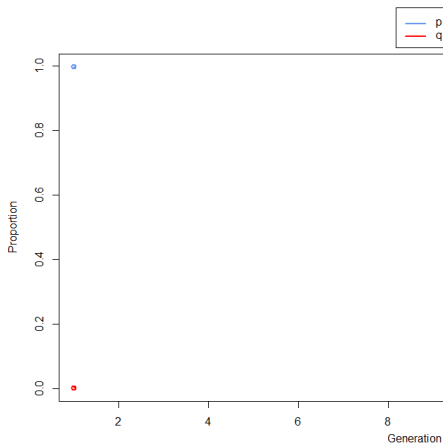


# Dynamics of a two Grammar system

$$p = 0.999 \quad q = 0.001 \quad \leftarrow \text{Actuation}$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_1 = \alpha p / (\alpha p + \beta q)$$

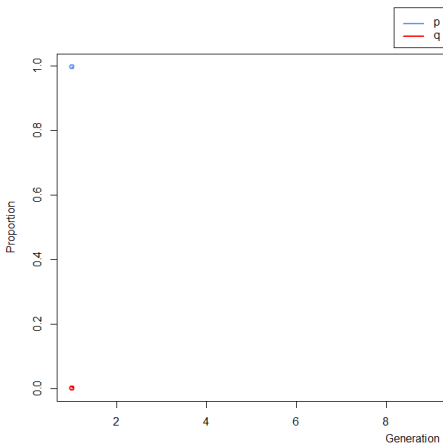


# Dynamics of a two Grammar system

$$p = 0.999 \quad q = 0.001$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_1 = 0.1 * 0.999 / (0.5 * 0.01 + 0.1 * 0.999) = 0.996$$

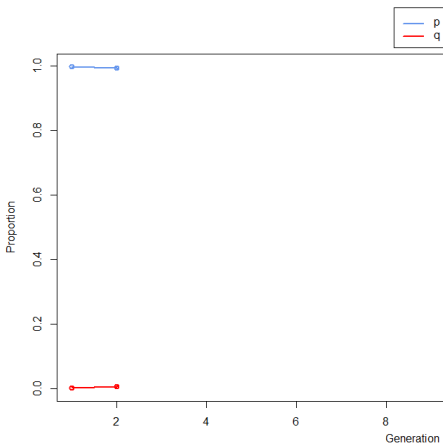


# Dynamics of a two Grammar system

$$p = 0.996 \quad q = 0.004$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_2 = 0.1 * 0.996 / (0.5 * 0.04 + 0.1 * 0.996) = 0.976$$





# Dynamics of a two Grammar system

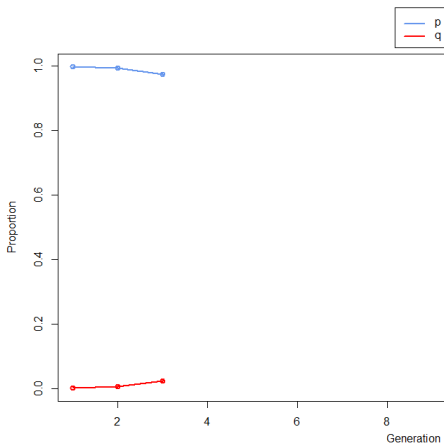
$$p = 0.976 \quad q = 0.024$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_3 = 0.1 * 0.976 /$$

$$(0.5 * 0.024 + 0.1 * 0.976) =$$

$$0.889$$

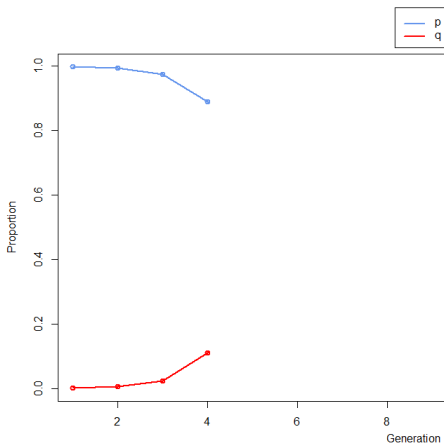


# Dynamics of a two Grammar system

$$p = 0.889 \quad q = 0.111$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_4 = 0.1 * 0.889 / \\ (0.5 * 0.111 + 0.1 * 0.889) = \\ 0.615$$

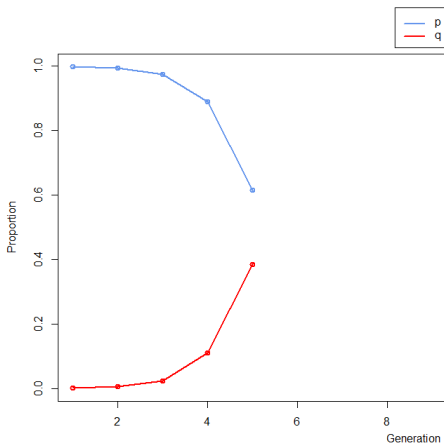


# Dynamics of a two Grammar system

$$p = 0.615 \quad q = 0.385$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_5 = 0.1 * 0.615 / (0.5 * 0.385 + 0.1 * 0.615) = 0.242$$

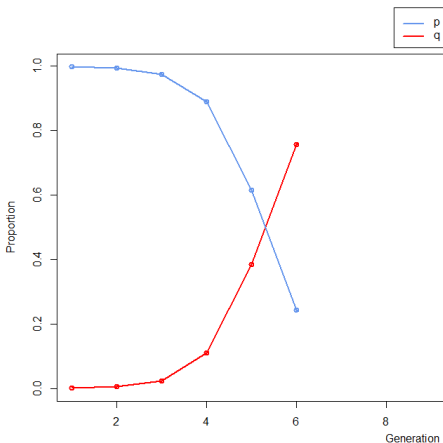


## Dynamics of a two Grammar system

$$p = 0.242 \quad q = 0.758$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_6 = 0.1 * 0.242 / \\ (0.5 * 0.758 + 0.1 * 0.242) = \\ 0.061$$

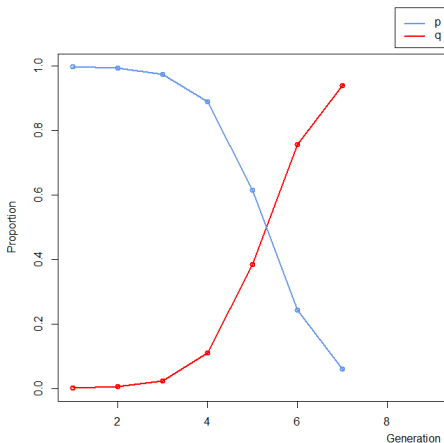


# Dynamics of a two Grammar system

$$p = 0.061 \quad q = 0.939$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_7 = 0.1 * 0.061 / (0.5 * 0.939 + 0.1 * 0.061) = 0.013$$



## Dynamics of a two Grammar system

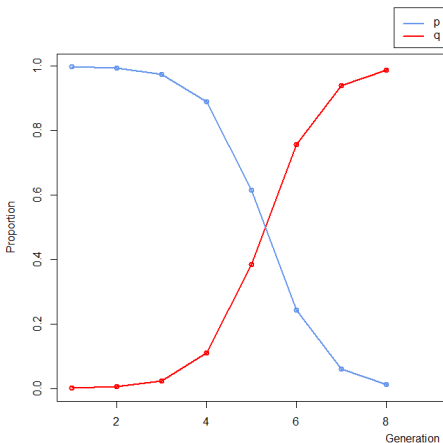
$$p = 0.013 \quad q = 0.978$$

$$\alpha = 0.1 \quad \beta = 0.5$$

$$p_8 = 0.1 * 0.013 /$$

$$(0.5 * 0.987 + 0.1 * 0.013) =$$

$$0.003$$

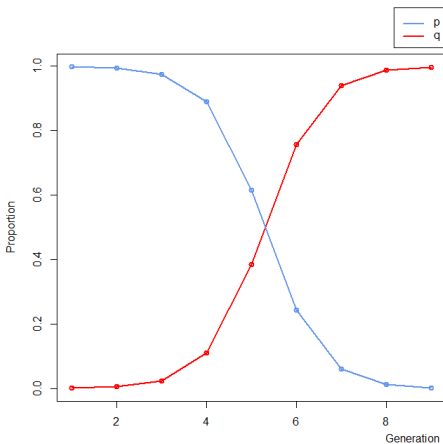


## Dynamics of a two Grammar system

$$p = 0.003 \quad q = 0.997$$

$$\alpha = 0.1 \quad \beta = 0.5$$

etc.



# Dynamics of a two Grammar system

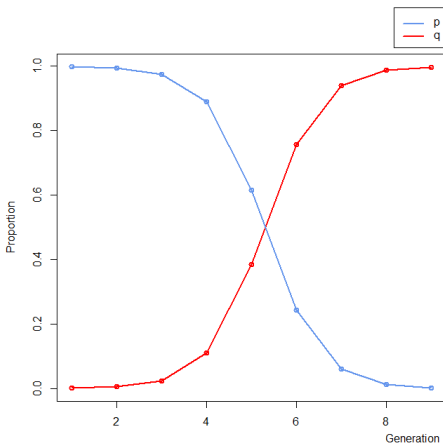
$$p = 0.003 \quad q = 0.997$$

$$\alpha = 0.1 \quad \beta = 0.5$$

*The fundamental theorem  
of language change:*

$G_2$  overtakes  $G_1$  if  $\beta > \alpha$

(Yang 2002, 239)





Yang's theory of language change

Development of early English IP headedness

Applying Yang's theory to the change in IP headedness

Outlook

Variation

Replicator dynamics

How to calculate fitness

# What is the advantage (or fitness) of a Grammar

## What is the advantage (or fitness) of a Grammar

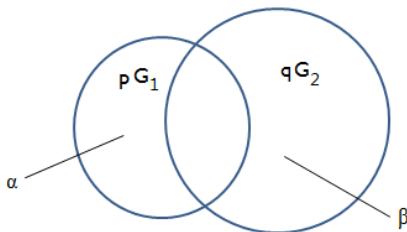
- ▶ "[A] proportion  $\alpha$  of  $G_1$  expressions are incompatible with  $G_2$ , and a proportion  $\beta$  of  $G_2$  expressions are incompatible with  $G_1$ . Call  $\alpha$  ( $\beta$ ) the *advantage* of  $G_1$  ( $G_2$ )." (Yang 2002, 130).

## What is the advantage (or fitness) of a Grammar

- ▶ "[A] proportion  $\alpha$  of  $G_1$  expressions are incompatible with  $G_2$ , and a proportion  $\beta$  of  $G_2$  expressions are incompatible with  $G_1$ . Call  $\alpha$  ( $\beta$ ) the *advantage* of  $G_1$  ( $G_2$ )." (Yang 2002, 130).
- ▶  $\alpha$  = proportion of sentences that can *only* be analyzed with  $G_1$   
 $\beta$  = proportion of sentences that can *only* be analyzed with  $G_2$

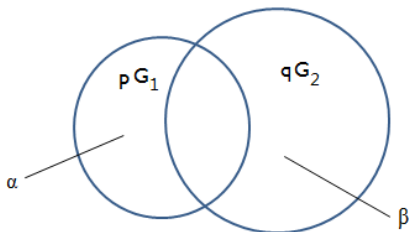
## What is the advantage (or fitness) of a Grammar

- ▶ "[A] proportion  $\alpha$  of  $G_1$  expressions are incompatible with  $G_2$ , and a proportion  $\beta$  of  $G_2$  expressions are incompatible with  $G_1$ . Call  $\alpha$  ( $\beta$ ) the *advantage* of  $G_1$  ( $G_2$ )." (Yang 2002, 130).
- ▶  $\alpha$  = proportion of sentences that can *only* be analyzed with  $G_1$   
 $\beta$  = proportion of sentences that can *only* be analyzed with  $G_2$



## What is the advantage (or fitness) of a Grammar

- ▶ "[A] proportion  $\alpha$  of  $G_1$  expressions are incompatible with  $G_2$ , and a proportion  $\beta$  of  $G_2$  expressions are incompatible with  $G_1$ . Call  $\alpha$  ( $\beta$ ) the *advantage* of  $G_1$  ( $G_2$ )." (Yang 2002, 130).
- ▶  $\alpha$  = proportion of sentences that can *only* be analyzed with  $G_1$   
 $\beta$  = proportion of sentences that can *only* be analyzed with  $G_2$



$\implies$  A less ambiguous Grammar is fitter (copying fidelity)

Yang's theory of language change

**Development of early English IP headedness**

Applying Yang's theory to the change in IP headedness

Outlook

The change

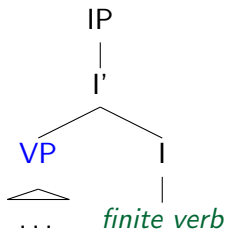
Illustration

# Development of early English IP headedness

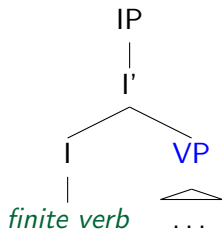
## From I-final to I-initial structure

- ▶ The finite verb comes to be placed more commonly clause-medially (Pintzuk 1999)
- ▶ Effect is strongest in subordinate clauses

I-final Grammar:



I-initial Grammar:



## IP headedness illustrated

(2) a. *I-final*

... þæt he [*VP* him wæpen] sealde

... that he        them weapons gave

'... that he gave them weapons'

(cobede, Bede\_2:10.138.2.1326)

b. *I-initial*

... þæt ic sealde [*VP* him Chanaan land]

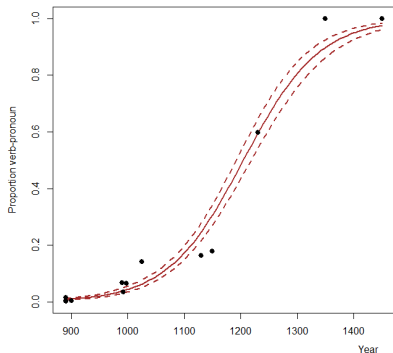
... that I gave        them Chanaan land

'...that I gave them the land of the Chanaanites'

(cootest, Exod:6.2.2527)



# Verb-pronoun order in some select texts



Text	Year	pro - verb	verb - pro
Bede	890	189	1
Cura Pastoralis	890	498	1
Orosius	890	255	4
Boethius	897	257	1
Cath. Hom. 1	990	322	23
Cath. Hom. 2	992	292	11
Lives of Saints	998	298	21
Apollonius	1025	30	5
Trinity Hom.	1130	82	16
Vices Virtues	1150	82	18
Ancrene Riwe	1230	87	129
Earliest Prose Ps.	1350	0	34
Margery Kempe	1450	0	100

York Corpus of Old English (Taylor et al. 2003), Penn Parsed Corpus of Middle English (Kroch and Taylor 2000)

Yang's theory of language change

Development of early English IP headedness

**Applying Yang's theory to the change in IP headedness**

Outlook

Step 1

Step 2

Step 3

# Applying Yang's theory to the change in IP headedness

Yang's theory of language change

Development of early English IP headedness

**Applying Yang's theory to the change in IP headedness**

Outlook

Step 1

Step 2

Step 3

# Road map

## Road map

- ▶ Step 1: Identify language stages that are purely  $G_1$  and  $G_2$  and that are measurable with corpora ("monolingual corpora")

## Road map

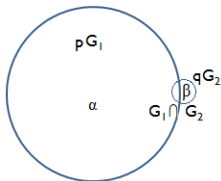
- ▶ Step 1: Identify language stages that are purely  $G_1$  and  $G_2$  and that are measurable with corpora ("monolingual corpora")
- ▶ Step 2: Estimate the advantage  $\alpha$  and  $\beta$  from these corpora (i.e. the proportion of expressions that would be incompatible with the "intruder" Grammar)

## Road map

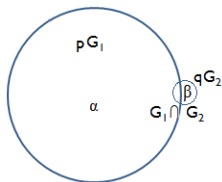
- ▶ Step 1: Identify language stages that are purely  $G_1$  and  $G_2$  and that are measurable with corpora ("monolingual corpora")
- ▶ Step 2: Estimate the advantage  $\alpha$  and  $\beta$  from these corpora (i.e. the proportion of expressions that would be incompatible with the "intruder" Grammar)
- ▶ Step 3: Determine if the Grammar that historically won did indeed have the larger advantage

(cf. Wallenberg 2013)

# Step 1: Identifying pure language stages



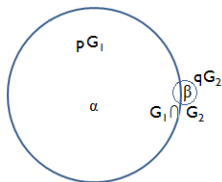
## Step 1: Identifying pure language stages



- ▶ There is no exclusively monolingual corpus with all I-final structures in the history of English. However, in the earliest periods, the new I-initial Grammar is vanishingly small: virtually no subordinate clauses are necessarily I-initial.

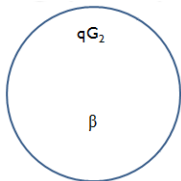


## Step 1: Identifying pure language stages

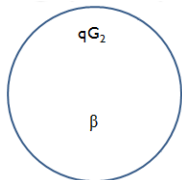


- ▶ There is no exclusively monolingual corpus with all I-final structures in the history of English. However, in the earliest periods, the new I-initial Grammar is vanishingly small: virtually no subordinate clauses are necessarily I-initial.
- ▶ Conservative Grammar  $G_1$  sample: subordinate clauses in *St. Chad, Letter of Alexander, Marvels of the East, Bede, Cura Pastoralis, Boethius*, earliest parts of the *Anglo-Saxon Chronicle*  
(York Corpus of Old English)

## Step 1: Identifying pure language stages

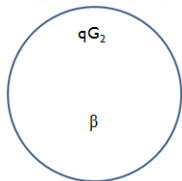


## Step 1: Identifying pure language stages



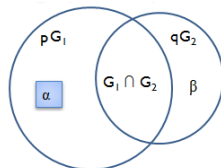
- ▶ Exclusively monolingual corpora with all I-initial structures are easy to find in the history of English: In Modern English virtually all sentences are generated by an I-initial Grammar.

## Step 1: Identifying pure language stages



- ▶ Exclusively monolingual corpora with all I-initial structures are easy to find in the history of English: In Modern English virtually all sentences are generated by an I-initial Grammar.
- ▶ Innovative Grammar  $G_2$  sample: subordinate clauses in all texts in periods M3 and M4 (1350-1500) (e.g., *Wycliffe Bible*, Chaucer's *Tale of Melibee*, Caxton's *Reynard the Fox*,...)  
(Penn Parsed Corpus of Middle English)

## Step 2: Estimating the advantages $\alpha$ and $\beta$

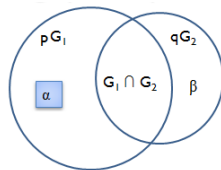


### (3) *pre-verbal heavy non-subject elements*

... þæt mon [þa stowe] [mid fyre] onæleð  
... that one the place with fire burned

'... so that the place was burned'  
(comarvel, Marv:6.6.35)

## Step 2: Estimating the advantages $\alpha$ and $\beta$



### (4) *nonfinite verb - finite verb*

... hu hi him wæpen [wyrcean] **sceoldan**  
... how they themselves weapons work should

'... how they should make arms for themselves'  
(cobede, Bede\_1:9.46.8.394)

## Step 2: Estimating the advantages $\alpha$ and $\beta$

► Productive postposition processes in early English

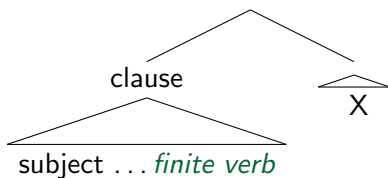
- (5) ... þæt he þam þegne **forgeafe** [bearnas wæstm].  
... that he the nobleman gave child's fruit  
' ... that he would give the child's fruit to the nobleman'  
(coeuphr,LS\_7\_[Euphr]:16.18)

## Step 2: Estimating the advantages $\alpha$ and $\beta$

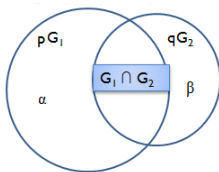
► Productive postposition processes in early English

- (6) ... hu þu [gehæled beon] **meaht** [from þisse aðle  
... how you healed be might from this disease  
hefignesse]  
heaviness  
'...how you might be healed from this grivious disease'  
(cobede, Bede\_3:10.186.22.1881)



Step 2: Estimating the advantages  $\alpha$  and  $\beta$ 

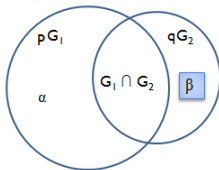
## Step 2: Estimating the advantages $\alpha$ and $\beta$



- ▶ As a consequence, every SV(X) clause will be ambiguous

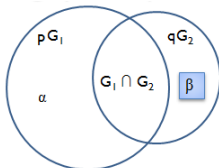
- (7) a. ... þæt hi **namon** [pæne þryddan dæl]  
... that they took the third part  
'... that they took the third part'  
(coelive, ÆLS-[Basil]:286.639)
- b. ... [they took [<sub>VP</sub> the third part]] (= I-initial)
- c. ... [they took] [the third part] (= potentially I-final)

## Step 2: Estimating the advantages $\alpha$ and $\beta$



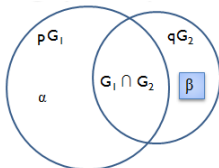
- (8) *postverbal non-postposing elements*
- a. ... if suddayne fortune **caused** **hyt.** pronoun  
(CMMALORY,650.4286)
  - b. ... whan þei **gon in** particle  
(CMMANDEV,55.1363)
  - c. ... þe walle þat he **stondeþ bi.** stranded preposition  
(CMHORSES,89.33)
  - d. ...

## Step 2: Estimating the advantages $\alpha$ and $\beta$



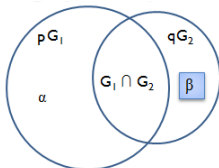
- ▶ postverbal diagnostic elements that cannot be analyzed with an I-final Grammar:

## Step 2: Estimating the advantages $\alpha$ and $\beta$



- ▶ postverbal diagnostic elements that cannot be analyzed with an I-final Grammar:
- ▶ non-subject personal pronouns, pronoun+*self*, bare demonstratives (*that, this*), negation and negative adverbs, negatively quantified objects, particles, directional adverbs (*forth, away*), stranded prepositions

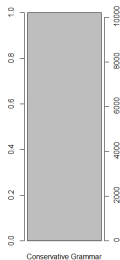
## Step 2: Estimating the advantages $\alpha$ and $\beta$



- ▶ postverbal diagnostic elements that cannot be analyzed with an I-final Grammar:
- ▶ non-subject personal pronouns, pronoun+*self*, bare demonstratives (*that, this*), negation and negative adverbs, negatively quantified objects, particles, directional adverbs (*forth, away*), stranded prepositions
- ▶ good evidence that at least some of these elements *never* postpose (e.g., Pintzuk and Haerberli 2008)

## Step 3: Comparing the advantages - $G_1$

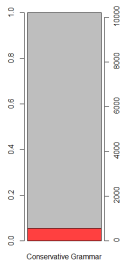
- ▶ All in all 10,981 subordinate clauses with an overt nominative subject in sample for conservative Grammar  $G_1$



## Step 3: Comparing the advantages - $G_1$

► 562 SOV clauses

- (9) ... þæt hie us [fersc wæter &  
... that they us fresh water and  
swete] **getæhton**.  
sweet showed  
'... so that they might show us fresh  
and sweet water'  
(coalex,Alex:15.5.132)

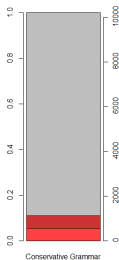




## Step 3: Comparing the advantages - $G_1$

- ▶ 565 SXV clauses (X=heavy phrase)

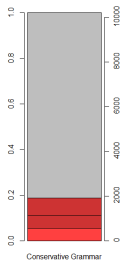
(10) ... hwelce hylde he [mid ðære  
... which favour he with the  
ælmessan] **gewriexle**  
alms gain  
'... which favour he might gain from  
the alms'  
(cocura,CP:44.323.17.2168).



## Step 3: Comparing the advantages - $G_1$

- ▶ 793 S nonfinite-verb V clauses

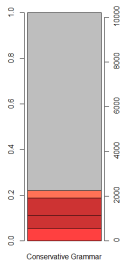
(11) ... hwæt þa þing [beon] sceolde  
... what that thing be should  
'... what that thing should be'  
(cobede, Bede\_4:3.264.30.2696)



## Step 3: Comparing the advantages - $G_1$

► 327 SOVX clauses

(12) ... þet þu hit [nenegum men] **ne**  
... that you it no man not  
**asecge** [er minre forðfore]  
say before my departure  
'... that you may not say it to any  
man before my departure'  
(cohad,LS\_3-[Chad]:134.90)

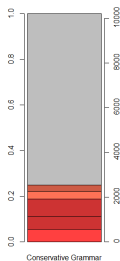


## Step 3: Comparing the advantages - $G_1$

- ▶ 298  $SX_1OX_2$  clauses ( $X_1 =$  heavy phrase)

(13) ... þæt hiera mægas [him mid]  
... that their relatives with him  
wæron [þa þe him from  
were those who him from  
noldon]  
not-would

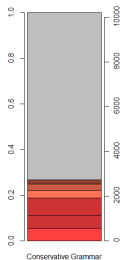
'... that their relatives were with him,  
who would not desert him'  
(ChronA\_[Plummer]:755.29.534)



## Step 3: Comparing the advantages - $G_1$

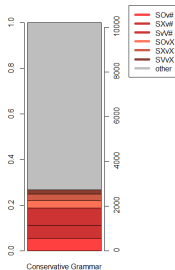
- ▶ 178 S nonfinite-verb VX clauses

(14) ... þæt ge lange [libban] **scylan**  
... that you long live shall  
[her on worulde]  
here in world  
'... that you shall live long here in  
this world'  
(coboeth,Bo:19.46.30.840)



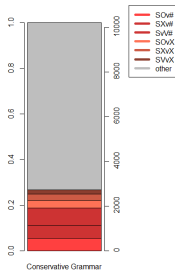
## Step 3: Comparing the advantages - $G_1$

- ▶ 2,723 clauses out of 10,198 clauses are unambiguously I-final



## Step 3: Comparing the advantages - $G_1$

- ▶ 2,723 clauses out of 10,198 clauses are unambiguously I-final
- ▶  $G_1$  has an advantage of  $\alpha = 27\%$



## Step 3: Comparing the advantages - $G_2$

- ▶ All in all 22,702 subordinate clauses with an overt subject (NP-SBJ) in sample for innovative Grammar  $G_2$

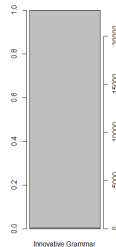




## Step 3: Comparing the advantages - $G_2$

- ▶ 82 clauses have post-verbal bare demonstratives

(15) ... when þe kyng and þe quene herde  
 þat  
 (CMBRUT3,20.573)



## Step 3: Comparing the advantages - $G_2$

- ▶ 114 clauses have relevant post-verbal directional adverbs

(16) ... alle thingis, to whiche he **zed** **forth**  
(CMPURVEY,I,10.366)



## Step 3: Comparing the advantages - $G_2$

- ▶ 467 clauses have post-verbal negation

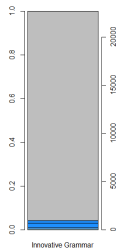
(17) ... that it **distourbith** **not** the  
instrument  
(CMASTRO,663.C1.46)



## Step 3: Comparing the advantages - $G_2$

- ▶ 276 clauses have post-verbal negatively quantified objects

(18) ... þat God **dop** **no wrong** to hem  
(CMWYCSE,382.2817)



## Step 3: Comparing the advantages - $G_2$

- ▶ 39 clauses have post-verbal pronouns with *self*

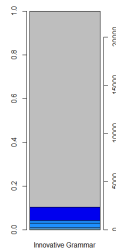
(19) ... as he *seip* *hymself* in þe gospel  
(CMVICES4,112.298)



## Step 3: Comparing the advantages - $G_2$

- ▶ 1353 clauses have post-verbal pronouns

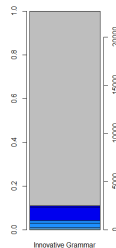
(20) ... as þis 3ong man **teld** **hym**  
(CMKEMPE,57.1271)



## Step 3: Comparing the advantages - $G_2$

- ▶ 158 clauses have post-verbal particles

(21) ... þat it may come out þe bettur  
(CMHORSES,119.327)

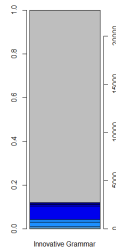


## Step 3: Comparing the advantages - $G_2$

- ▶ 251 clauses have post-verbal stranded prepositions

(22) ... the worlde that I **have** moste  
desyre **to**.

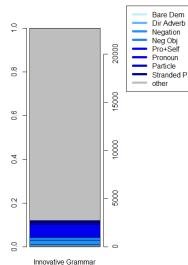
(CMMALORY,669.4949)





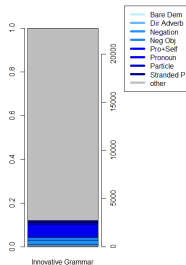
## Step 3: Comparing the advantages - $G_2$

- ▶ 2,740 clauses out of 22,702 clauses are unambiguously I-initial

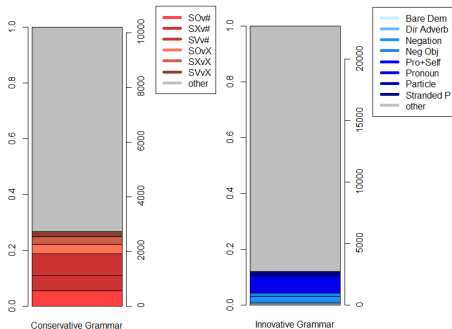


## Step 3: Comparing the advantages - $G_2$

- ▶ 2,740 clauses out of 22,702 clauses are unambiguously I-initial
- ▶  $G_2$  has an advantage of  $\beta = 12\%$



## Step 3: Comparing the advantages



$$\alpha \approx 0.27$$

$$\beta \approx 0.12$$

⇒ Hypothesis that advantage  $\beta$  of I-initial Grammar is greater than advantage  $\alpha$  of I-final Grammar could *not* be confirmed

## Evaluation

- ▶ On the face of it, the case study counts as a falsification of Yang's theory of language change

## Evaluation

- ▶ On the face of it, the case study counts as a falsification of Yang's theory of language change
- ▶ However, there are potential reasons for the negative result:

# Evaluation

- ▶ On the face of it, the case study counts as a falsification of Yang's theory of language change
- ▶ However, there are potential reasons for the negative result:
  1. Unambiguous contexts for  $G_2$  missing? In particular, SVO sentences are perhaps more likely to be interpreted as evidence for  $G_2$  as postposition declines?

# Evaluation

- ▶ On the face of it, the case study counts as a falsification of Yang's theory of language change
- ▶ However, there are potential reasons for the negative result:
  1. Unambiguous contexts for  $G_2$  missing? In particular, SVO sentences are perhaps more likely to be interpreted as evidence for  $G_2$  as postposition declines?
  2. Unambiguous contexts for  $G_1$  over-interpreted? In particular, SXV sentences and sentences with postposition may not really count as evidence for  $G_1$ .

## Evaluation

- ▶ On the face of it, the case study counts as a falsification of Yang's theory of language change
- ▶ However, there are potential reasons for the negative result:
  1. Unambiguous contexts for  $G_2$  missing? In particular, SVO sentences are perhaps more likely to be interpreted as evidence for  $G_2$  as postposition declines?
  2. Unambiguous contexts for  $G_1$  over-interpreted? In particular, SXV sentences and sentences with postposition may not really count as evidence for  $G_1$ .
- ▶ Therefore, Yang may still be right.



## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.

## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.
- ▶ ⊖ Advantage as 'copying fidelity' probably incomplete:

## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.
- ▶ ⊖ Advantage as 'copying fidelity' probably incomplete:
  1. Simplicity, economy considerations and other formal (innate, grammar-specific) preferences (e.g., Clark and Roberts 1993).

## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.
- ▶ ⊖ Advantage as 'copying fidelity' probably incomplete:
  1. Simplicity, economy considerations and other formal (innate, grammar-specific) preferences (e.g., Clark and Roberts 1993).
  2. Processing ease, discourse processing and other functional (general, cognitive) preferences (e.g., Hawkins 1994).

## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.
- ▶ ⊖ Advantage as 'copying fidelity' probably incomplete:
  1. Simplicity, economy considerations and other formal (innate, grammar-specific) preferences (e.g., Clark and Roberts 1993).
  2. Processing ease, discourse processing and other functional (general, cognitive) preferences (e.g., Hawkins 1994).
  3. Analogical processes. E.g. I-initial structures in Old English main clauses; prevalence of *do*-support for negation before extension to possessive *have* in American English.

## My view on Yang's theory of language change (1/3)

- ▶ ⊕ Fitness - language change as selection and not random drift.
- ▶ ⊖ Advantage as 'copying fidelity' probably incomplete:
  1. Simplicity, economy considerations and other formal (innate, grammar-specific) preferences (e.g., Clark and Roberts 1993).
  2. Processing ease, discourse processing and other functional (general, cognitive) preferences (e.g., Hawkins 1994).
  3. Analogical processes. E.g. I-initial structures in Old English main clauses; prevalence of *do*-support for negation before extension to possessive *have* in American English.
  4. Sociological pressures. Syntax can be a more or less prestigious identity marker. E.g. habitual *be* in AAVE (e.g., Escalles 1994).

## My view on Yang's theory of language change (2/3)

- ▶ ⊕ Empirically testable in principle with corpora. Predictive.

## My view on Yang's theory of language change (2/3)

- ▶ ⊕ Empirically testable in principle with corpora. Predictive.
- ▶ ⊖ Testing is not a trivial, maybe even unfeasible task:



## My view on Yang's theory of language change (2/3)

- ▶ ⊕ Empirically testable in principle with corpora. Predictive.
- ▶ ⊖ Testing is not a trivial, maybe even unfeasible task:
  1. There is no standardized falsification procedure for Yang's claims. Measuring of advantages must be **objective** (but researchers know the outcome of a change beforehand), **reliable** (but no two language states are ever identical) and **valid** (but the advantage measured depends heavily on specific structural assumptions).

## My view on Yang's theory of language change (2/3)

- ▶ ⊕ Empirically testable in principle with corpora. Predictive.
- ▶ ⊖ Testing is not a trivial, maybe even unfeasible task:
  1. There is no standardized falsification procedure for Yang's claims. Measuring of advantages must be **objective** (but researchers know the outcome of a change beforehand), **reliable** (but no two language states are ever identical) and **valid** (but the advantage measured depends heavily on specific structural assumptions).
  2. Dubious representativeness of currently available corpora.

## My view on Yang's theory of language change (2/3)

- ▶ ⊕ Empirically testable in principle with corpora. Predictive.
- ▶ ⊖ Testing is not a trivial, maybe even unfeasible task:
  1. There is no standardized falsification procedure for Yang's claims. Measuring of advantages must be **objective** (but researchers know the outcome of a change beforehand), **reliable** (but no two language states are ever identical) and **valid** (but the advantage measured depends heavily on specific structural assumptions).
  2. Dubious representativeness of currently available corpora.
  3. Complexity of language system. Even few lexical items may potentially cause a large-scale change. But not every aspect that may be relevant for advantages, let alone every lexical item, can be scrutinized.

## My view on Yang's theory of language change (3/3)

- ▶ ⊕ Mathematical rigid. Simple. Usable for linguists.

## My view on Yang's theory of language change (3/3)

- ▶ ⊕ Mathematical rigid. Simple. Usable for linguists.
- ▶ ⊖ Yang's mathematical model is not without alternatives.  
Could the same ideas be formulated more elegantly?

## My view on Yang's theory of language change (3/3)

- ▶ ⊕ Mathematical rigid. Simple. Usable for linguists.
- ▶ ⊖ Yang's mathematical model is not without alternatives.  
Could the same ideas be formulated more elegantly?
  1. Good mathematical tools in historical linguistics are multivariate statistics, in particular logistic regression models, which have great descriptive and predictive power. Discrete generations, at  $t$ ,  $t+1$ ,  $t+2$  etc., may not be as elegant as a continuous time variable. The goodness of fit of observed data to a logistic regression model can be assessed (e.g. residual analysis). Yang does not explain how the goodness of fit of observed data should be assessed for his model.

## My view on Yang's theory of language change (3/3)

- ▶ ⊕ Mathematical rigid. Simple. Usable for linguists.
- ▶ ⊖ Yang's mathematical model is not without alternatives. Could the same ideas be formulated more elegantly?
  1. Good mathematical tools in historical linguistics are multivariate statistics, in particular logistic regression models, which have great descriptive and predictive power. Discrete generations, at  $t$ ,  $t+1$ ,  $t+2$  etc., may not be as elegant as a continuous time variable. The goodness of fit of observed data to a logistic regression model can be assessed (e.g. residual analysis). Yang does not explain how the goodness of fit of observed data should be assessed for his model.
  2. Evolutionary game theory offers a very attractive mathematical way to think about replicating populations (e.g., Jäger 2008). Using game theory may save us from reinventing the mathematical wheel.

Thank you for your attention!

$$\dot{x}(t) = -x(t) \left[ 2u_x x(t) - \Gamma \int_{t_0}^t dt' G_y(t, t') x(t') - \kappa_x(t) + \eta_x(t) \right]$$

$$\dot{y}(t) = -y(t) \left[ 2u_y y(t) - \Gamma \int_{t_0}^t dt' G_x(t, t') y(t') - \kappa_y(t) + \eta_y(t) \right]$$

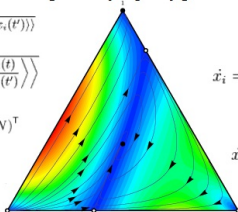
Yang's theory is pretty great!

$$C_x(t, t') = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \overline{\langle \langle x_i(t) x_i(t') \rangle \rangle}$$

$$G_x(t, t') = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \overline{\langle \langle \frac{\delta x_i(t)}{\delta \kappa_x(t')} \rangle \rangle}$$

$$\dot{N}_i(t) = N_i(t) e_i \cdot W(\sigma(t), N)^T$$

$$\dot{N} = Na \left( 1 - \frac{N}{K} \right)$$



$$\dot{x}_i = \sum_{j=1}^n x_j f_j(x) Q_{ji} - \phi(x) x_i$$

$$\dot{x}_i = x_i \left( (Ax)_i - x^T Ax \right)$$

Pah, just stick with  
logistic regression.

Maybe I should use  
Game Theory instead...

$$\dot{x}_i = x_i [f_i(x) - \phi(x)], \quad \phi(x) = \sum_{i=1}^n x_i f_i(x)$$

$$Z[\psi^x, \psi^y] = \left\langle \left\langle \exp \left( i \int dt \left\{ \sum_{i=1}^N \psi_i^x(t) x_i(t) + \sum_{j=1}^N \psi_j^y(t) y_j(t) \right\} \right) \right\rangle \right\rangle$$



- Clark, R. and Roberts, I. (1993), 'a computational model of language learnability and language change', *Linguistic Inquiry* **24**, 299–345.
- Croft, W., ed. (2000), *Explaining Language Change: An Evolutionary Approach*, Longman Linguistics Library, London.
- Davies, M. (2010), *The Corpus of Historical American English: 400 million words, 1810-2009*, <http://corpus.byu.edu/coha/> (Accessed 10 September 2013).
- Escales, J. E. (1994), 'African american vernacular english in advertising: a sociolinguistic study', *Advances in Consumer Research* **21**, 304–309.
- Hawkins, J. A., ed. (1994), *A Performance Theory of Order and Constituency*, Cambridge University Press, Cambridge.
- Jäger, G. (2008), 'Applications of game theory in linguistics', *Language and Linguistics Compass* **2**, 408–421.

Kroch, A. (1989), 'Reflexes of grammar in patterns of language change', *Journal of Language Variation and Change* **1.3**, 199–244.

Kroch, A. and Taylor, A. (2000), *Penn-Helsinki Parsed Corpus of Middle English*,  
<http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3>  
(Accessed 10 April 2013), 2 edn, Department of Linguistics,  
University of Pennsylvania.

Pintzuk, S. (1999), *Phrase Structures in Competition: Variation and Change in Old English Word Order*, Garland, New York.

Pintzuk, S. and Haeberli, E. (2008), 'Structural variation in old english root clauses', *Language Variation and Change* **20**, 367–407.

- Taylor, A., Warner, A., Pintzuk, S. and Beths, F. (2003), *The York-Toronto-Helsinki Corpus of Old English Prose (YCOE)*, <http://www-users.york.ac.uk/lang22/YCOE/YcoeHome.htm> (Accessed 10 April 2013), Oxford Text Archive.
- Wallenberg, J. (2013), 'Scrambling, If, and phrase structure change in yiddish', *Lingua* **133**, 289–318.
- Yang, C. (2002), *Knowledge and Learning in Natural Language*, Oxford University Press, Oxford.