Aging and Language Change

Ellis Cain

Abstract

Language change can be understood as representing the dynamic relationship between group-level and individual usage patterns. According to statistical learning, language learners are sensitive to the underlying distributional properties of language usage. Therefore, when usage patterns vary across time, it would be expected that the language acquisition of individuals will be impacted by the usage patterns that they experience during their lifetime. While previous research has shown differences in the language representation and cognitive capabilities over the lifespan (i.e., early vocab spurt, late cognitive decline), the impact of varying learning rates on language change is not well understood. Using an agent-based model of (syntactic) language change, this project aims to explore the interaction between learning rates and aging may impact diachronic changes in usage patterns of language.

Introduction

Language change

Verb bias example

Aging

Previous models of language change

Beeksma et al., 2017

Troutman et al., ?

Model

Overview

I aim to explore how potential changes in learning rates across the lifespan may impact overall population-level patterns of language change. The model will explore the usage and spread of a grammatical variant throughout a population. For example, the past tense ending has some variation, either appearing as “-t” or “-ed”, such as in “learnt” or “learned”. [**verb bias example**] *To define: learn / learning rates* The following main assumptions are based on those found in Beeksma et al. and Troutman et al.:

1. Language learning is based on imitating others, though this may change over the lifespan.
    - E.g., individuals may learn quickly early on, but slow down as they age.

2. There are variations in preference between individuals.
    - E.g., some individuals learn more quickly than others.
3. Language can be influenced by external factors.
    - E.g., more willing to learn from in-group members, natural bias towards a given variant.

With this model, the main outcome patterns are as follows (adapted from Troutman et al.):

1. S-shaped curve in usage patterns: Change happens slowly, then proceeds rapidly before slowing down again.
2. Intra-speaker variation: Change is gradual and there is a period of intra-speaker variation.
3. Categorical norms: With competition, speakers move toward categorically using just one of the competing variants.
4. Multi-stability: Language change can have multiple stable outcomes. **Additionally, some variation in language usage may co-exist. (Verb bias example)**
5. Threshold problem: Initially rare variants may manage to spread through entire speech communities.

Model description

This model is based on the language change model from Troutman et al., implemented in netlogo.

The model consists of a preferential attachment network, where each node represents a language user. There are two grammar variants, which are coded as 0 or 1.

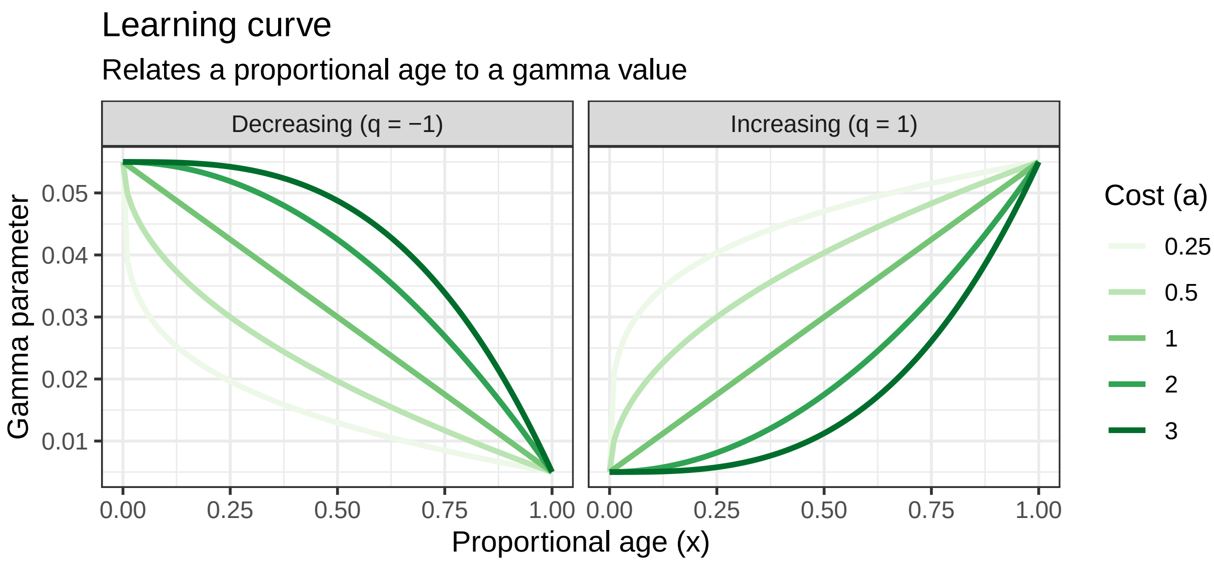
Nodes have the following properties; state, original state, spoken state, age, gamma, and cohort. State represents the node's current grammar state, ranging from [0,1], which will influence their likelihood to use one of the two grammar variants. It can be understood as the bias towards a specific grammar (e.g., node with state of 0.25 would be biased towards Grammar 0). Original state represents their initially assigned grammar state. Spoken state represents the output of each node's speech, as (grammar) 0 or 1. The speaking process will be detailed in the *dynamics* section. Each node also has an age, that increases by 1 each tick, and a gamma value, which serves as their learning rate (step size). Higher gamma values (0.04) mean that the node will adjust their state more when listening, as opposed to a lower gamma (0.01). Each node also has a cohort or age group, which is either 1 or 2. This is fixed from the start.

For initialization, the model uses [**what implementation**] to generate the preferential attachment network with a specified number of nodes. **Comment from Troutman paper?** Grammar is distributed throughout the network based on the specified percentage of grammar 1.

Next, the gamma, age, and cohort attributes of the nodes are initialized. If gamma is deterministic, the gamma values for all of the nodes will be set to the specified value. If gamma is probabilistic, the gamma values will be generated using a normal distribution centered around the initial gamma value, with a specified standard deviation. If gamma is set to be based on age, it will be determined by the "learning curve". The learning curve describes the (power law) relationship between age and learning rate, and is as follows:

$y = qc(x^a) + b$

where q indicates whether the learning rate increases ($+1$) or decreases (-1) with age (x). The constant c shrinks the function to the proper range, and is set to $c = 0.05$. The variable x is the normalized age of a given node, such that it is within [0,1]. The variable a represents the "cost" of learning, such that higher "costs" ($a = 3$) indicate slow initial change which increases quickly when x is closer to 1, and lower "costs" ($a = 0.25$) indicate a quick initial change that slows down later on. b is the initial value for gamma. Graphs for the increasing and decreasing learning curves can be seen below:



The number of cohorts is decided by specifying the percentage of Group 2. The cohorts are initialized such that their ages are either deterministic (set to the specified age) or probabilistic (centered around the specified age, with a specified standard deviation). When gamma is based on age, any variation in gamma will be determined by the setting of age distribution (deterministic or probabilistic), since they are generated from the learning curve.

Dynamics

Go procedure:

- Communication
  - Alpha parameter as bias in favor of grammar 1
- Aging
  - Age increases by 1 each tick
  - Gamma is modified by the specified parameters

Gamma modification:

- Influence is specified at the start: increases, decreases, remains constant
- Every 100 ticks, the gamma is modified by a specified constant, based on the direction
- Minimum gamma is 0; potential for some agents to stop learning

Communication:

- Speaking
  - Original model had variation in whether or not the agents preferred a discrete grammar; this model will take that as granted
  - Use a logistic curve, such that agents will produce an utterance that is either 0 or 1.

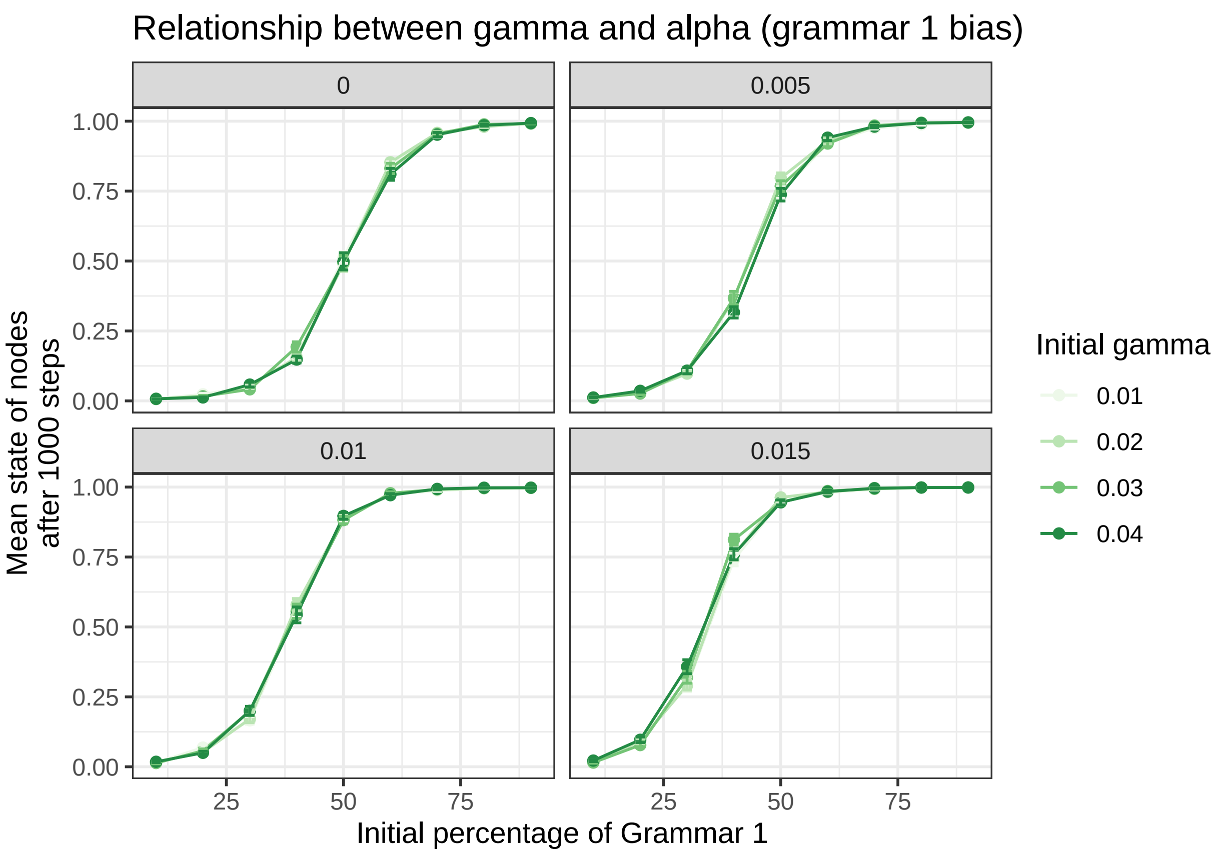
- Listening
  - Original model had variation in the listening function, either threshold (if neighbors above a threshold value, switch to that grammar), individual (select one neighbor, choose that grammar), or reward (explained below)
    - This model will take the third, reward-based algorithm as granted
  - Hearing node will pick a grammar that will be used to interpret utterances, either 0 or 1
  - If the selected grammar matches the heard grammar, it will update its grammar towards the heard state
  - If it fails, it will be updated away from the heard state

Results

Influence of gamma with only one group:

Low vs high gamma:

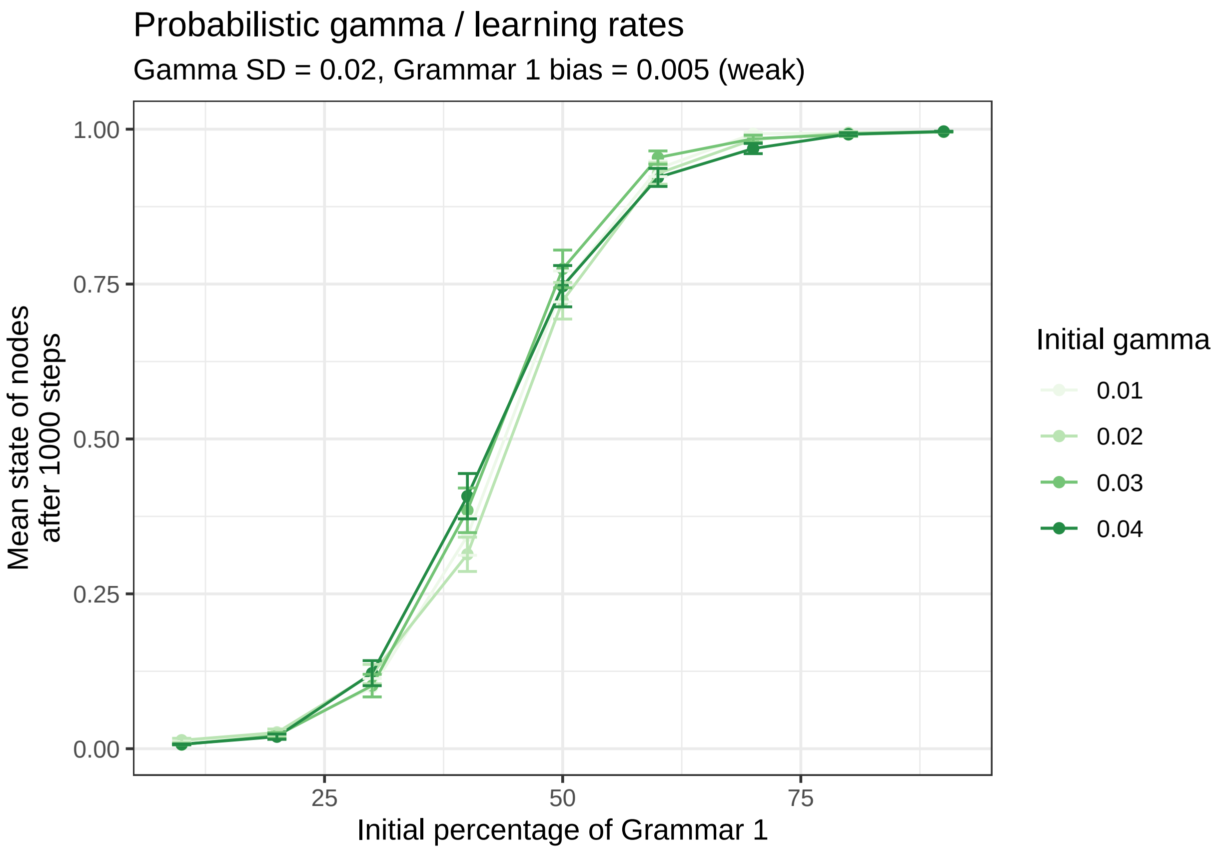
Use this style of summary graph



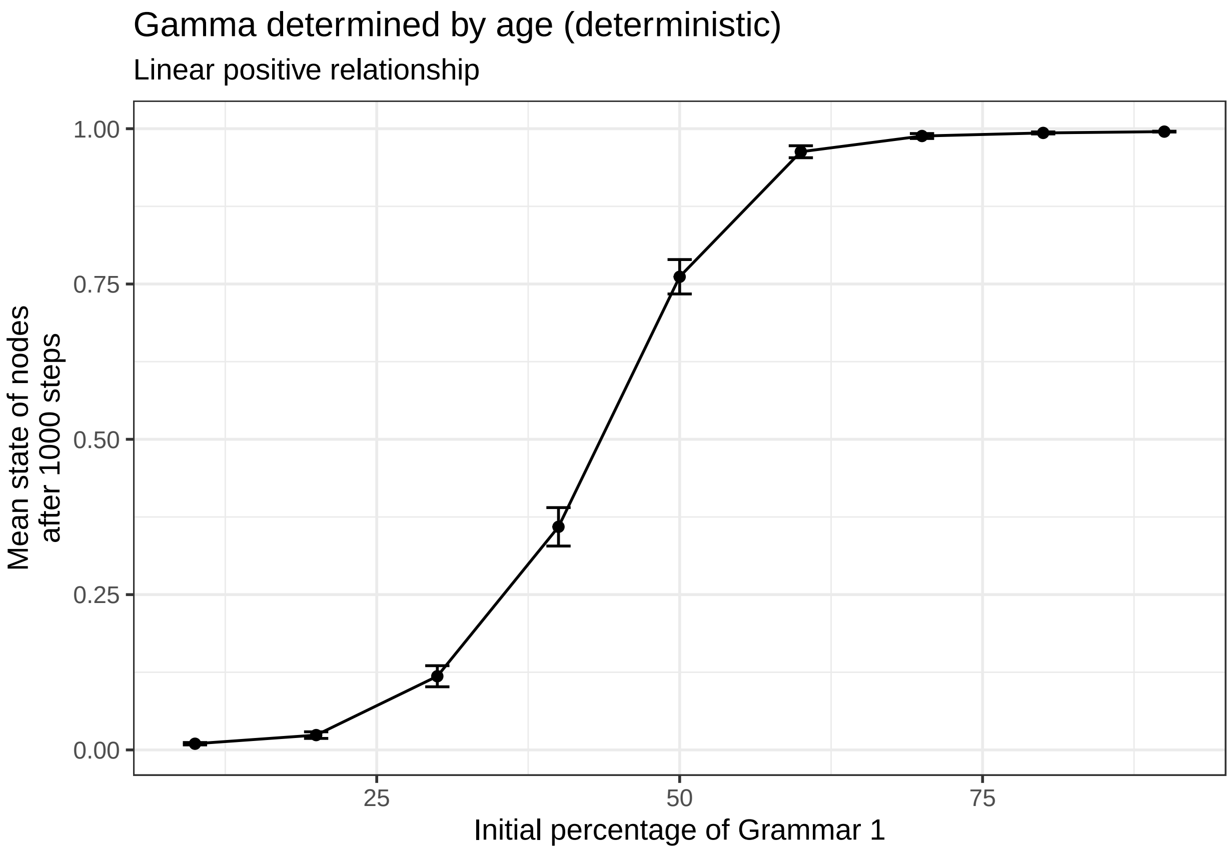
Individual runs:

```
## [1] "holder"
```

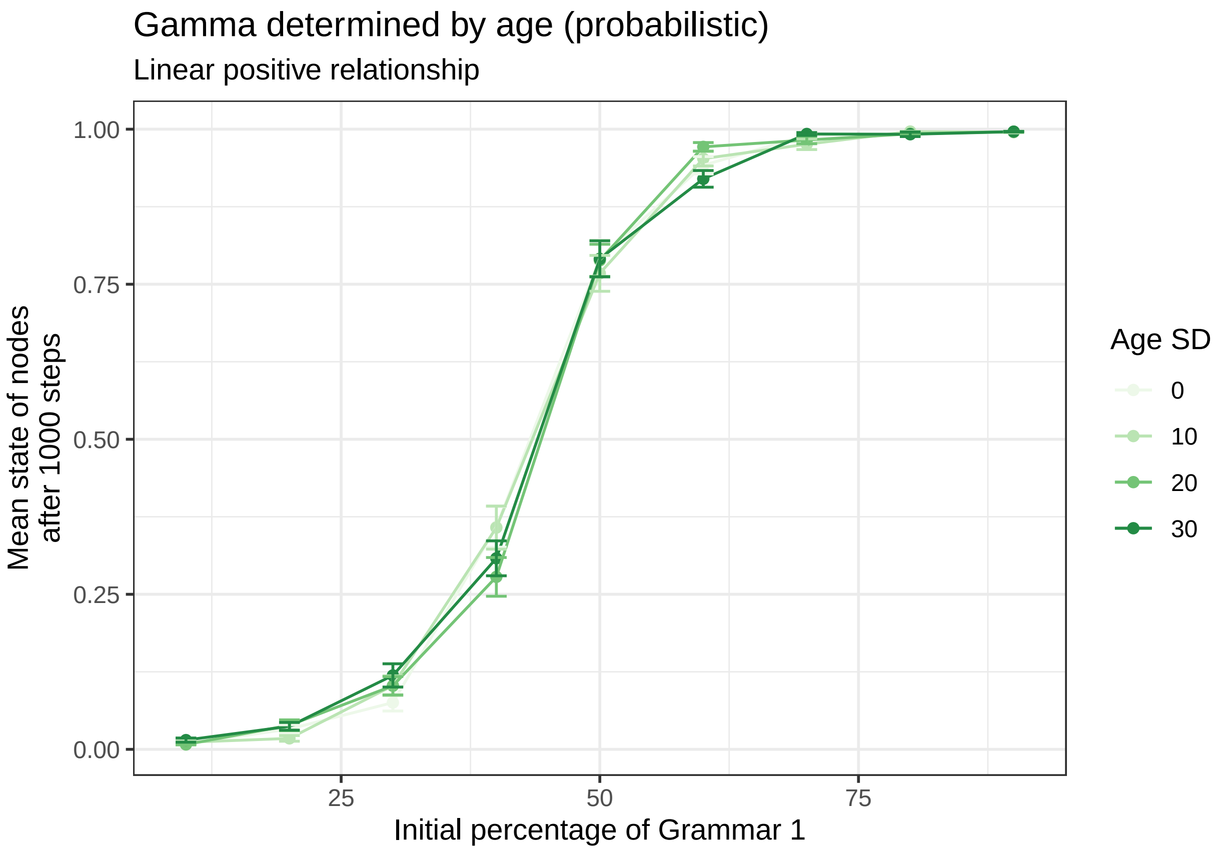
Probabilistic gamma:

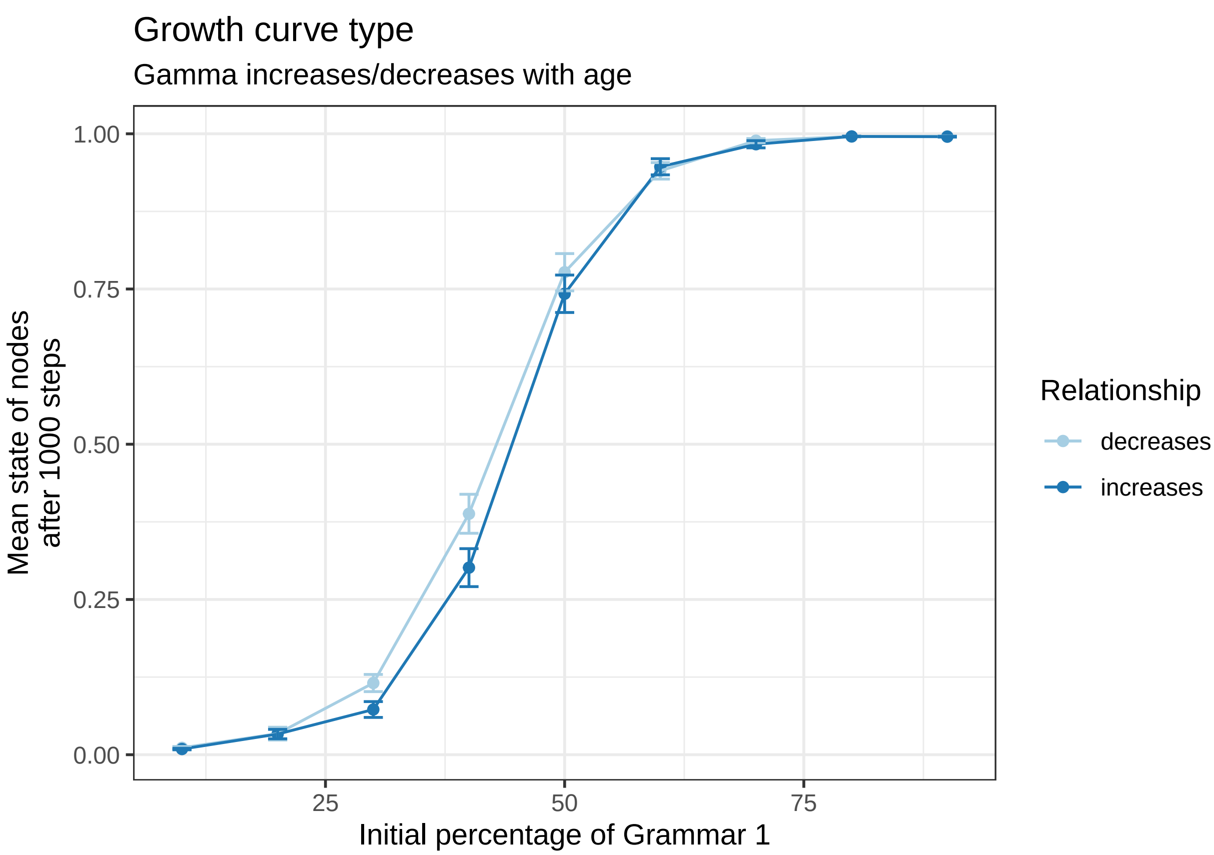


Aging and gamma:

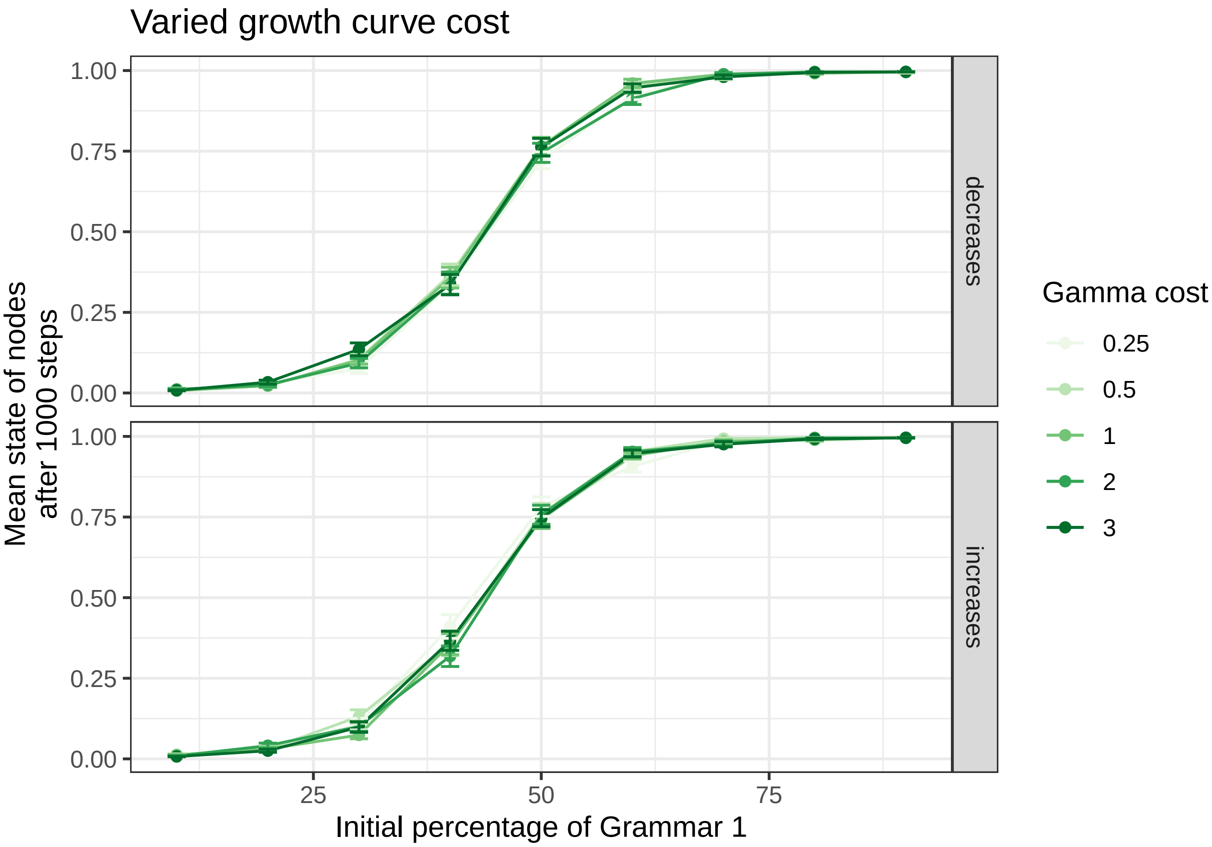

Gamma determined by age (deterministic)
Linear positive relationship
1.00
0.75
0.50
0.25
0.00
Mean state of nodes after 1000 steps
25
50
75
Initial percentage of Grammar 1

Age varies:

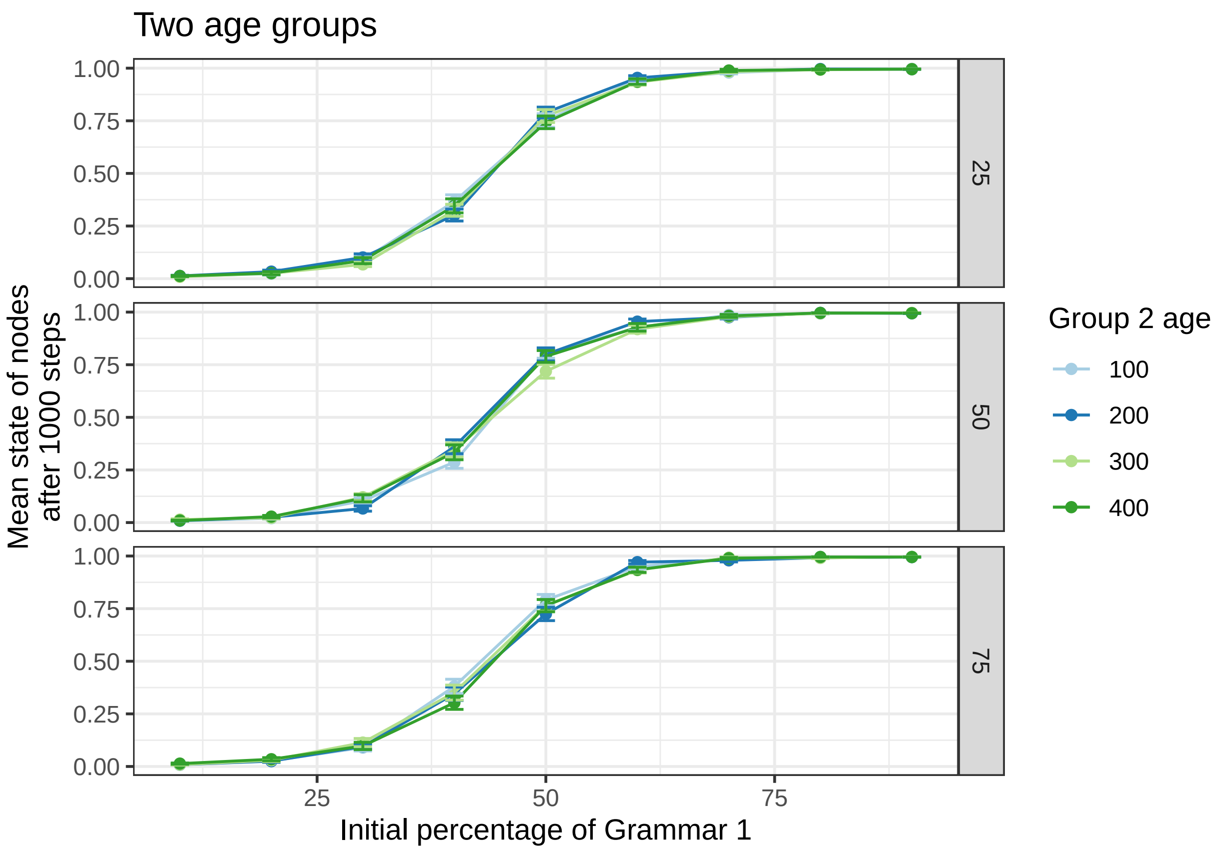

Gamma determined by age (probabilistic)
Linear positive relationship
Mean state of nodes after 1000 steps
1.00
0.75
0.50
0.25
0.00
25
50
75
Initial percentage of Grammar 1
Age SD
0
10
20
30


Growth curve type
Gamma increases/decreases with age
Mean state of nodes after 1000 steps
1.00
0.75
0.50
0.25
0.00
25
50
75
Initial percentage of Grammar 1
Relationship
decreases
increases

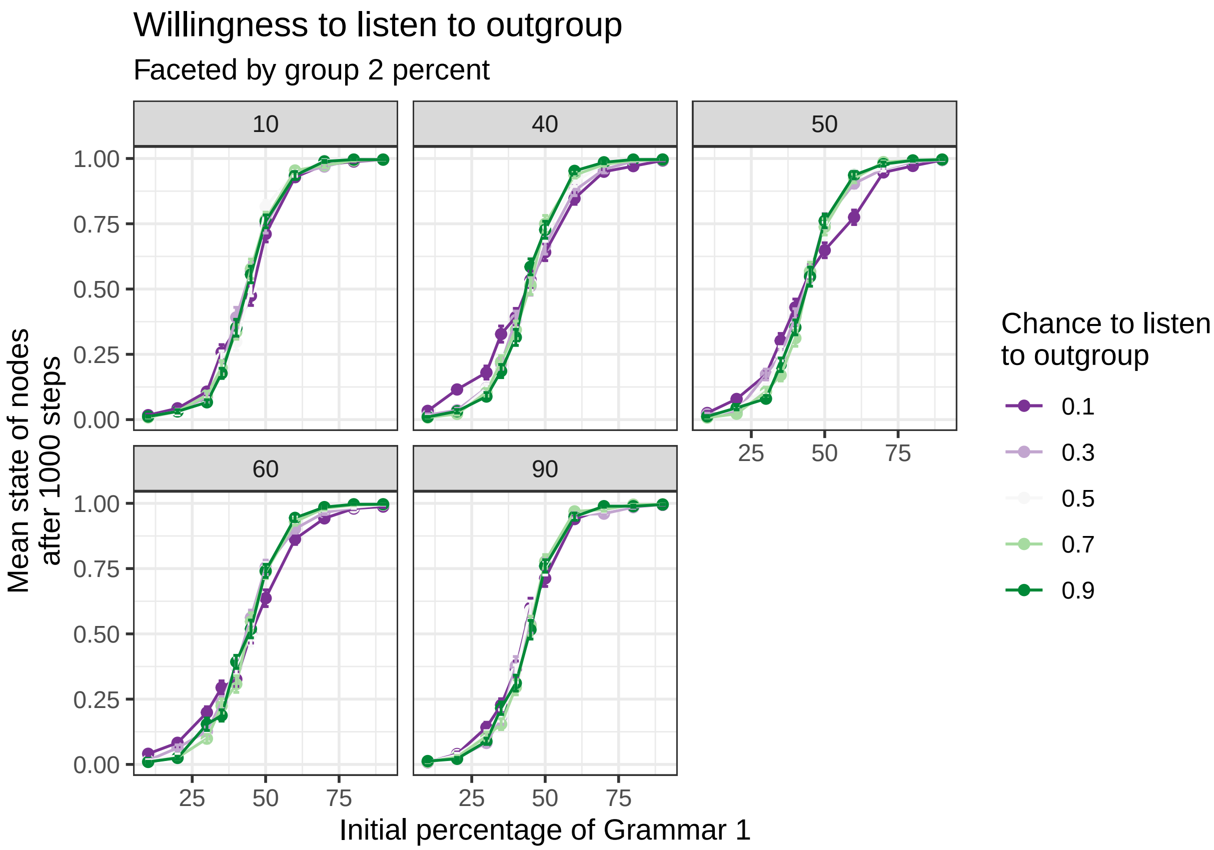
Varied cost curve



Influence of two groups:



Group preference:



Measure: average steps to equilibrium? Emergence of dialect subgroups?

Discussion

Summarize the results, in the context of previous literature

Comparison with previous model results

Life-long learning, language change

Conclusion.