

A stochastic model of language change through prediction-driven instability

W. Garrett Mitchener

College of Charleston

July 2007

Table of Contents

Introduction

Language variation & change

An unstructured population

Deterministic ODE

Markov chain

An age-structured population

Markov chain

Deterministic ODE

Two parameters

Independent

Dependent

Conclusion

X

Language variation

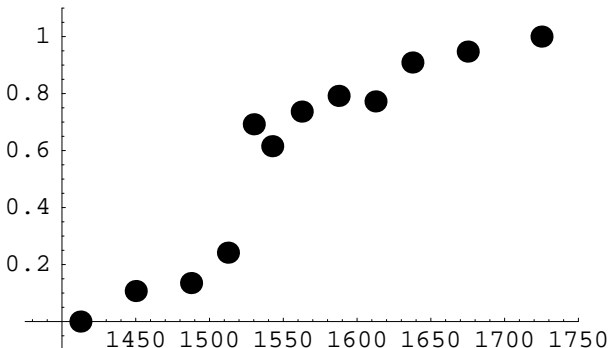
- Multiple ways to say the same thing
Ex: Late Middle English
 - (1) I know not the muffin man
 - (2) I do not know the muffin man
- Can be correlated to...
 - Location
 - Context
 - Gender
 - Attitude

Language change

Goal: Model spontaneous internally-driven language change

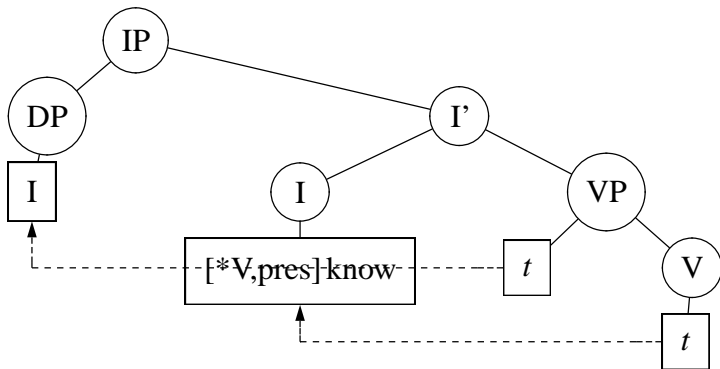
- Grows from variation
- Each speaker uses a mixture of old & new
- One variant replaces the other, mostly monotonically

Ex: *do*-support

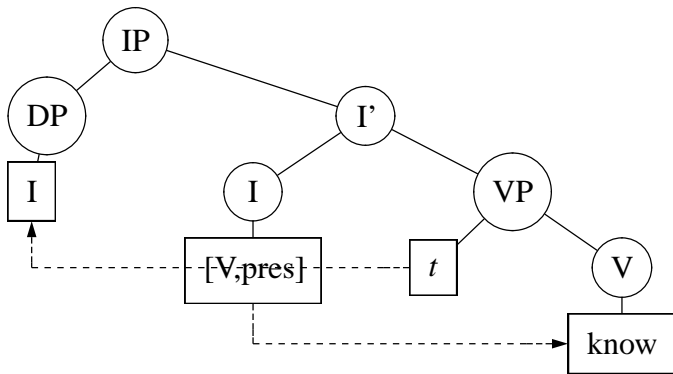


Transitive affirmative questions, fraction using *do* [Ellegård, 1953, Kroch, 1989, Warner, 2005]

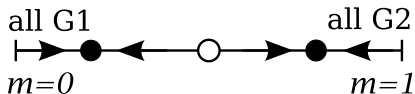
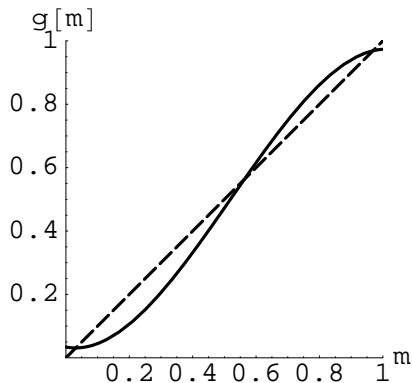
Verb raising



Affix hopping



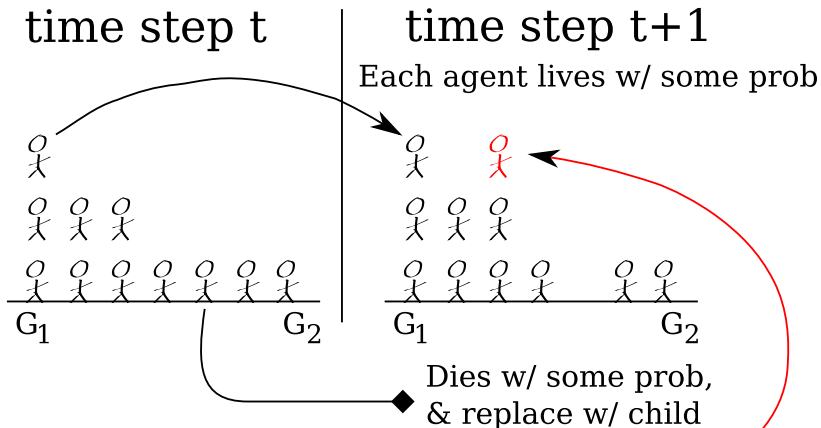
Mean field differential equation



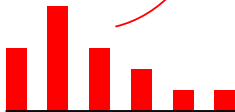
- One parameter \implies 2 grammars
 - G_1 : param unset
 - G_2 : param set
- m = mean rate of G_2
- $g(m)$ = mean rate of children
- Birth/death process

$$\frac{dm}{dt} = g(m) - m$$

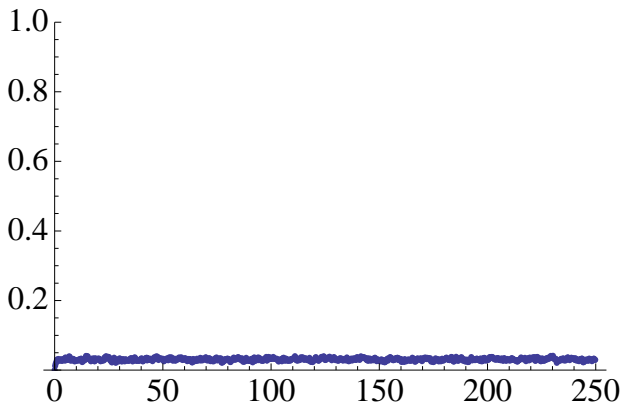
Unstructured Markov chain



Dist of children for time $t+1$
= fn of state at time t

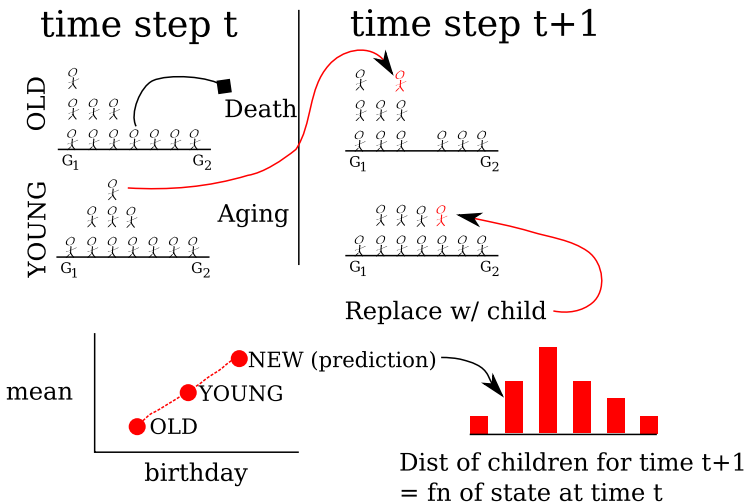


Time trace of one trajectory

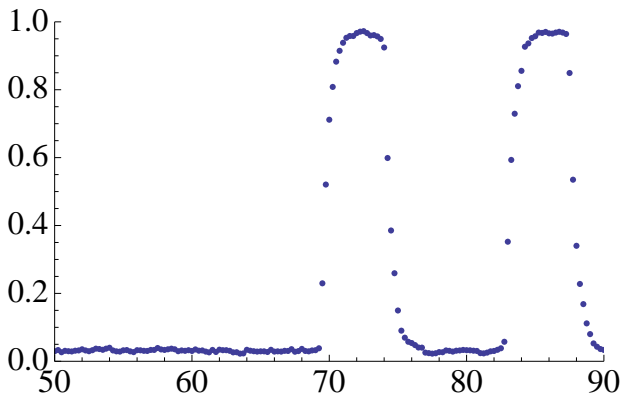


Mean usage rate of G_2 as a function of time

Age structured Markov chain: Incrementation

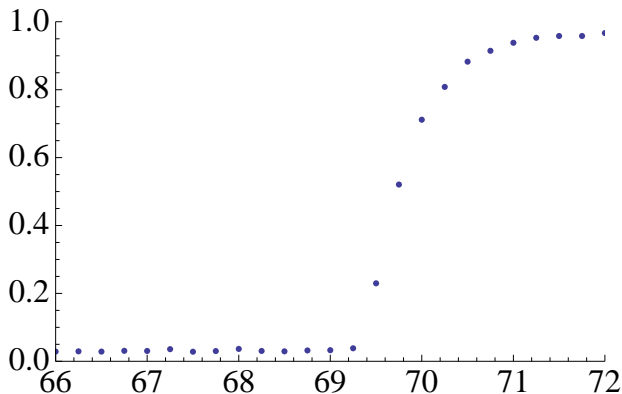


Time trace of one trajectory



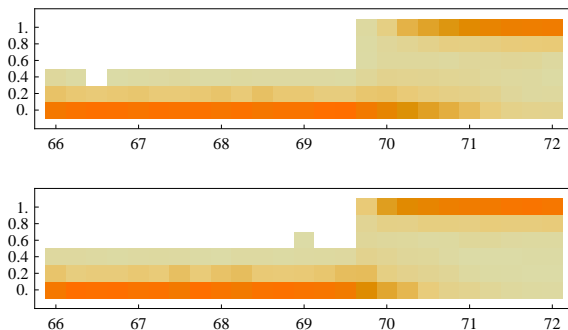
Mean usage rate of G_2 among young group as a function of time

Time trace of one trajectory: One transition



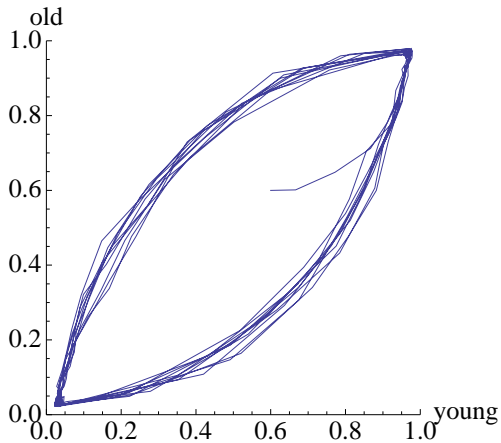
Mean usage rate of G_2 among young group as a function of time

Young distribution and old distribution

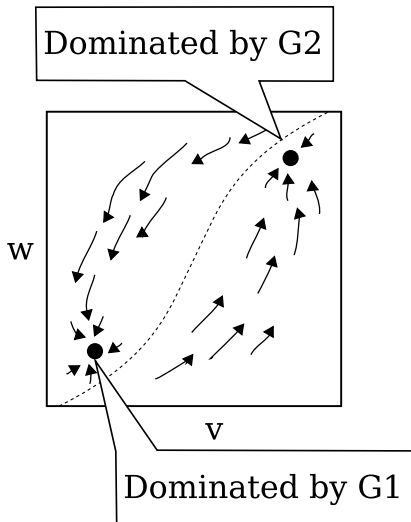


Number of people with usage rate z (vertical axes). Time increases left to right. Top: Old. Bottom: Young. Darker = more.

Old mean rate vs. young mean rate



Mean field differential equation w/ age structure



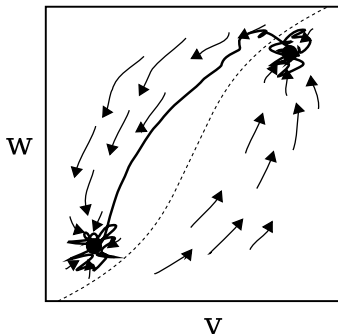
- v = mean usage rate of G_2 in the young group
- w = mean usage rate of G_2 in the old group
- Birth, learning, aging, death

$$\frac{dw}{dt} = v - w$$

$$\frac{dv}{dt} = g(r(v, w)) - v$$

- $r(v, w)$ = prediction

Mean field differential equation w/ age structure



- v = mean usage rate of G_2 in the young group
- w = mean usage rate of G_2 in the old group
- Birth, learning, aging, death

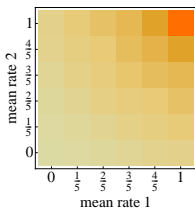
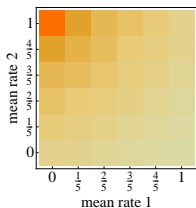
$$\frac{dw}{dt} = v - w$$

$$\frac{dv}{dt} = g(r(v, w)) - v$$

- $r(v, w)$ = prediction

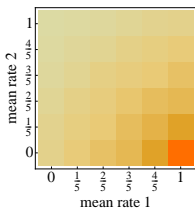
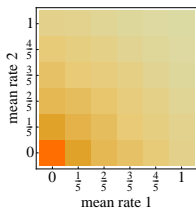
Learning two independent parameters

$$m_2 = \frac{31}{32}$$



- Parameters 1 & 2
- Means m_1 & m_2

$$m_2 = \frac{1}{32}$$

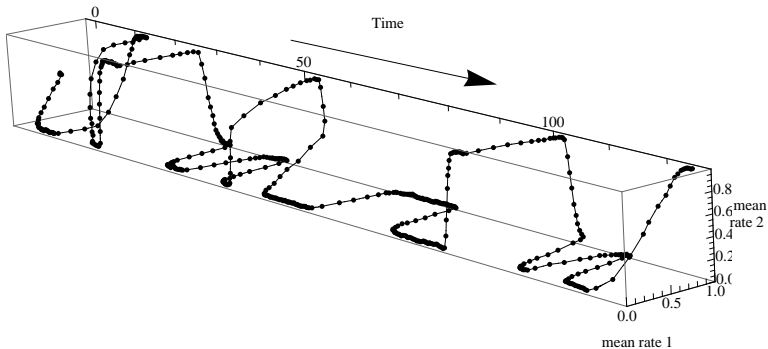


- Distribution of child speech = function of m_1 and m_2
- Darker = more

$$m_1 = \frac{1}{32}$$

$$m_1 = \frac{31}{32}$$

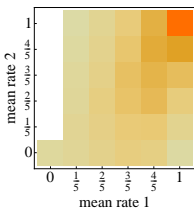
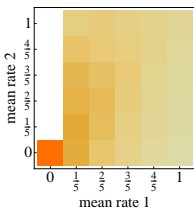
Two independent parameters: Time trace



Mean usage rates among young group as a function of time

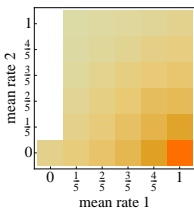
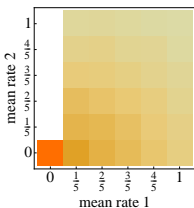
Learning two dependent parameters

$$m_2 = \frac{31}{32}$$



- Parameters 1 & 2
- Can only set 2 if 1 is set

$$m_2 = \frac{1}{32}$$

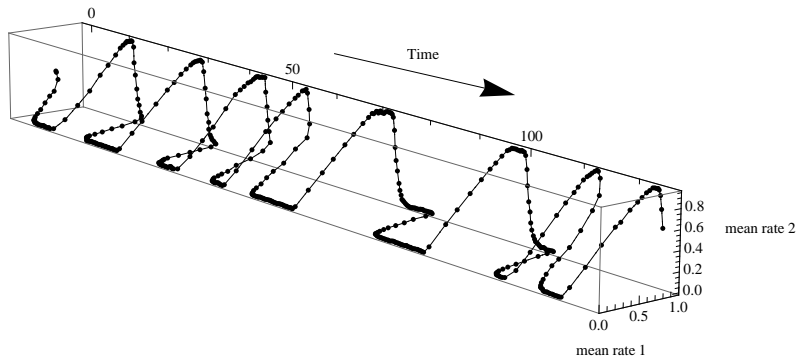


- Means m_1 & m_2
- Distribution of child speech = function of m_1 and m_2
- Darker = more

$$m_1 = \frac{1}{32}$$

$$m_1 = \frac{31}{32}$$

Two dependent parameters: Time trace



Mean usage rates among young group as a function of time

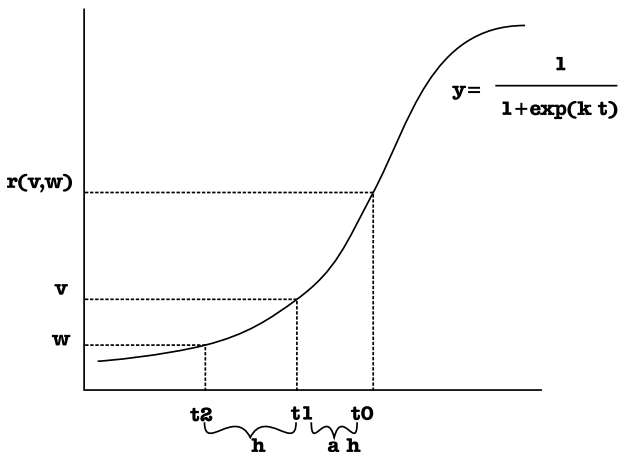
Conclusion

- Model of variable speech — usage rates
- In unstructured population language doesn't change spontaneously
- In structured population language can change spontaneously
- Prediction-driven instability

Bibliography

- Alvar Ellegård. *The Auxiliary do: The Establishment and Regulation of Its Use in English*. Gothenburg Studies in English. Almqvist and Wiksell, 1953.
- Anthony Kroch. Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1:199–244, 1989.
- Anthony Warner. Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change*, 17:257–280, 2005. DOI: 10.1017/S0954394505050106.

Prediction function



More *do*-support data

Do Support Frequencies

