Actuation without production bias

James Kirby, University of Edinburgh
<j.kirby@ed.ac.uk>

Morgan Sonderegger, McGill University
<morgan.sonderegger@mcgill.ca>

Sound Change in Interacting Human Systems
UC-Berkeley
May 31, 2014

Introduction

- Change at the population level is often claimed to be based in phonetic variation at the individual level (e.g. Ohala, 1993)
- One source of variation: production bias (e.g., coarticulation)

WGmc	*Pre-OHG*	*OHG (NHD)*
*gasti	gesti	gest (*Gäste)*
*lambir	lembir	lemb (*Lämme)*
*fasti	festi	fest (*fest)*

Primary umlaut in West Germanic (after Iverson and Salmons, 2006).

- This conference: other types of bias (group membership, cognitive endowment...)

Stability and change

- Existence of a bias does not mean change is inevitable: default is stability! (Weinrich et al., 1968; cf. Kiparsky's "non-phonologization problem")
- "Accumulation-of-error" approaches often criticized for this very reason (e.g. Baker, 2008)
- Adequate account of actuation must explain:
  1. Stability of limited coarticulation in the population;
  2. Stability of full coarticulation in the population;
  3. Change from stable limited to full coarticulation.

Roadmap

- First: summary of previous work showing that one way to get both stability and change at the population level is to assume both
  1. a force promoting contrast maintenance, to keep separate phonetic categories stable; and,
  2. an external force, such as a production bias, which induces change (cf. Pierrehumbert, 2001; Wedel, 2006).

Kirby & Sonderegger (2013), *Proc. CogSci*

Roadmap

- Then: today's questions
  1. Does using production bias as the external force have a unique dynamics?
  2. If not, will *any* kind of external force produce the same behaviour at the population level?
  3. Broader Q: can we safely assume that any proposed source of change *could* lead to change, iterated over time in a population?

Roadmap

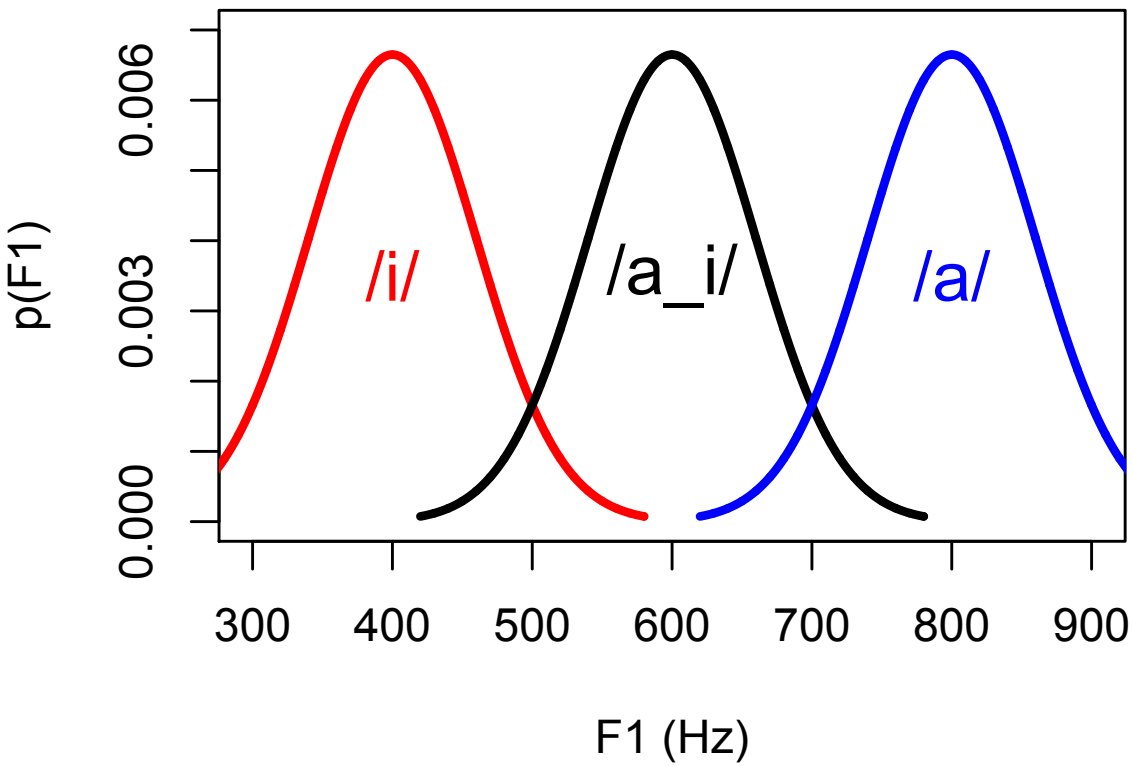
- Our example scenario: phonologization of coarticulation

WGmc	*Pre-OHG*	*OHG (NHD)*
*gasti	gesti	gest (*Gäste)*
*lambir	lembir	lemb (*Lämme)*
*fasti	festi	fest (*fest)*

- Simple models $\Rightarrow$ potentially unintuitive outcomes

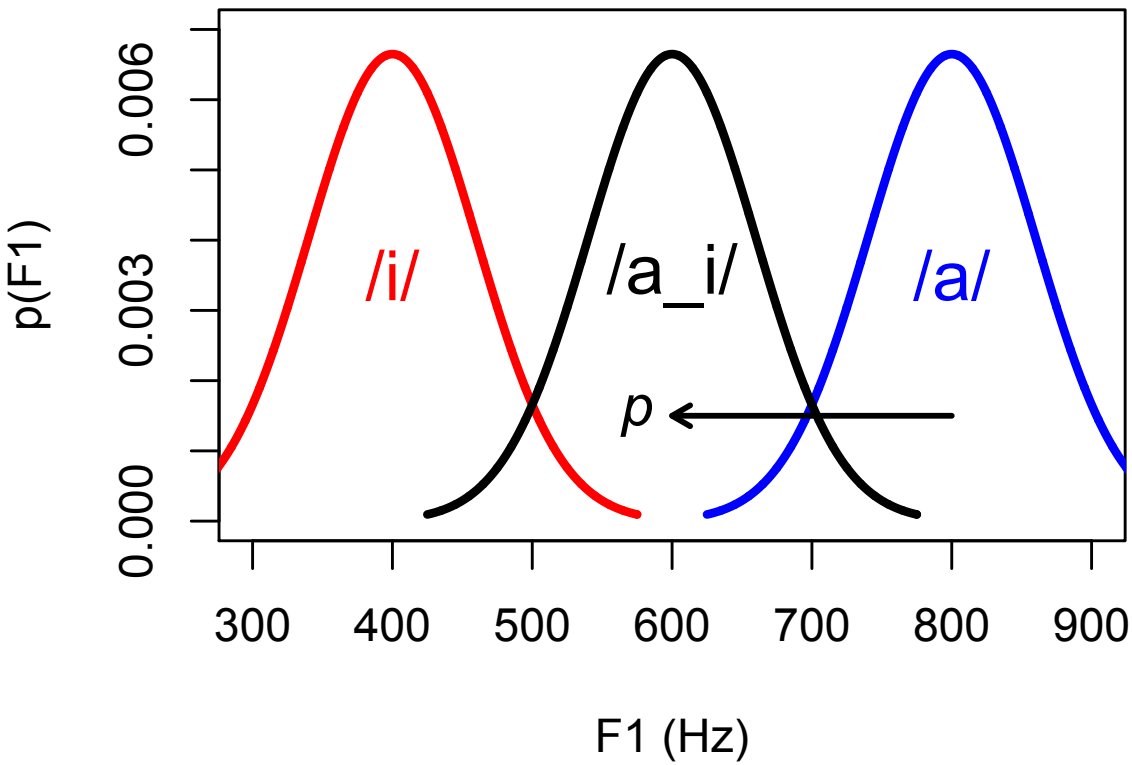
Framework

- *Lexicon*: $\{V_1, V_2, V_{12}\}$, where V_{12} represents V_1 in the coarticulation-inducing context of V_2



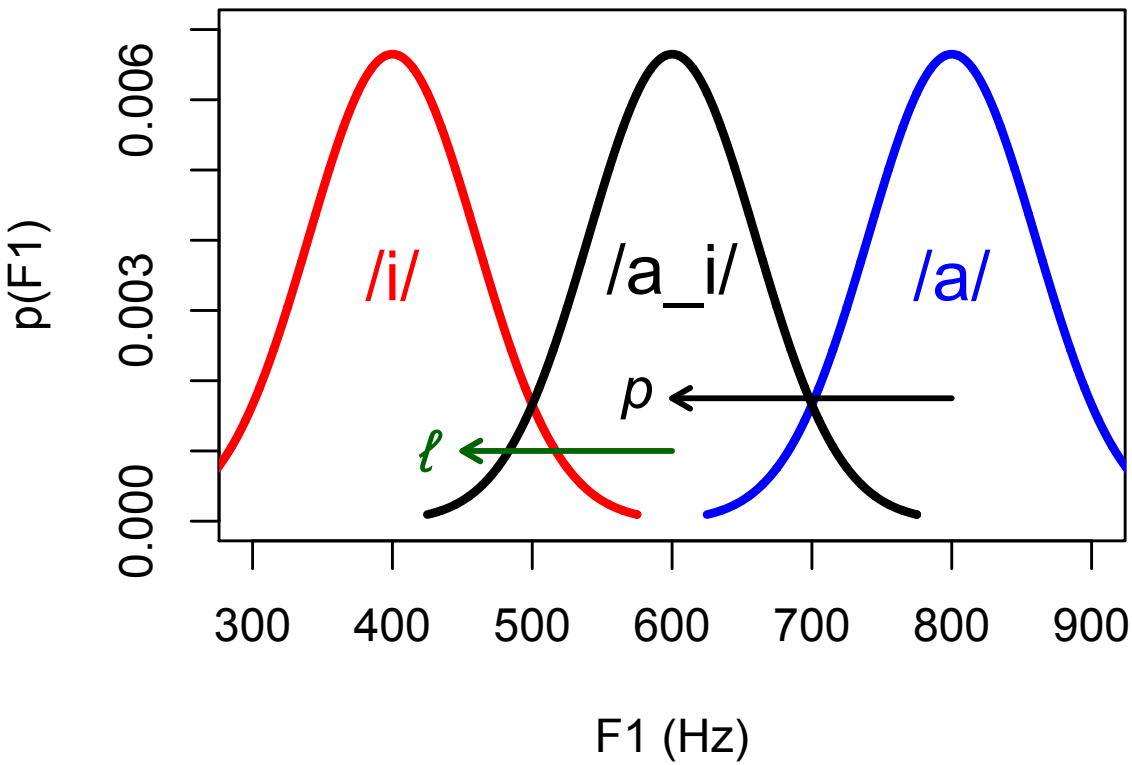
Framework

- *Task:* learn an **offset parameter** *p*: how much /a/ is produced like /i/ in the context of /i/ (/a_i/)



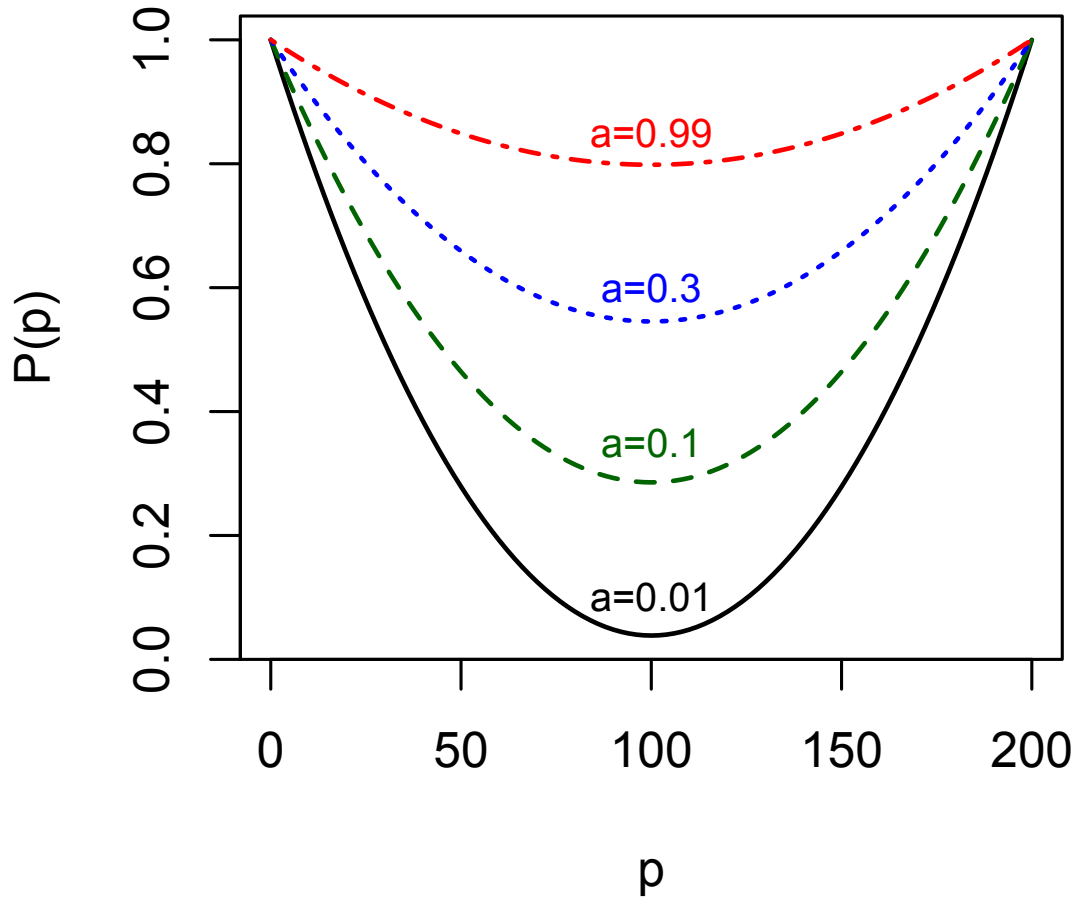
Framework

- *Data:* F1 values for /a_i/ tokens, potentially subject to **production bias** ℓ (assuming fixed /a/, /i/)



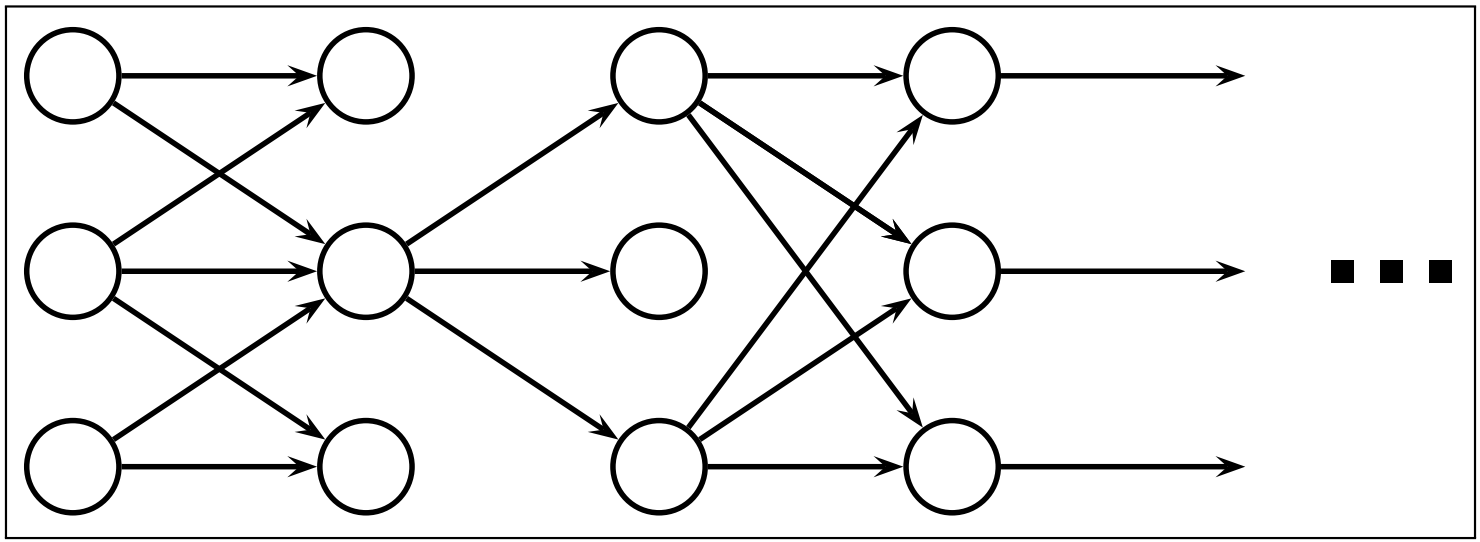
Framework

- *Learner's prior:* (strength of) **categoricity bias** (CB)



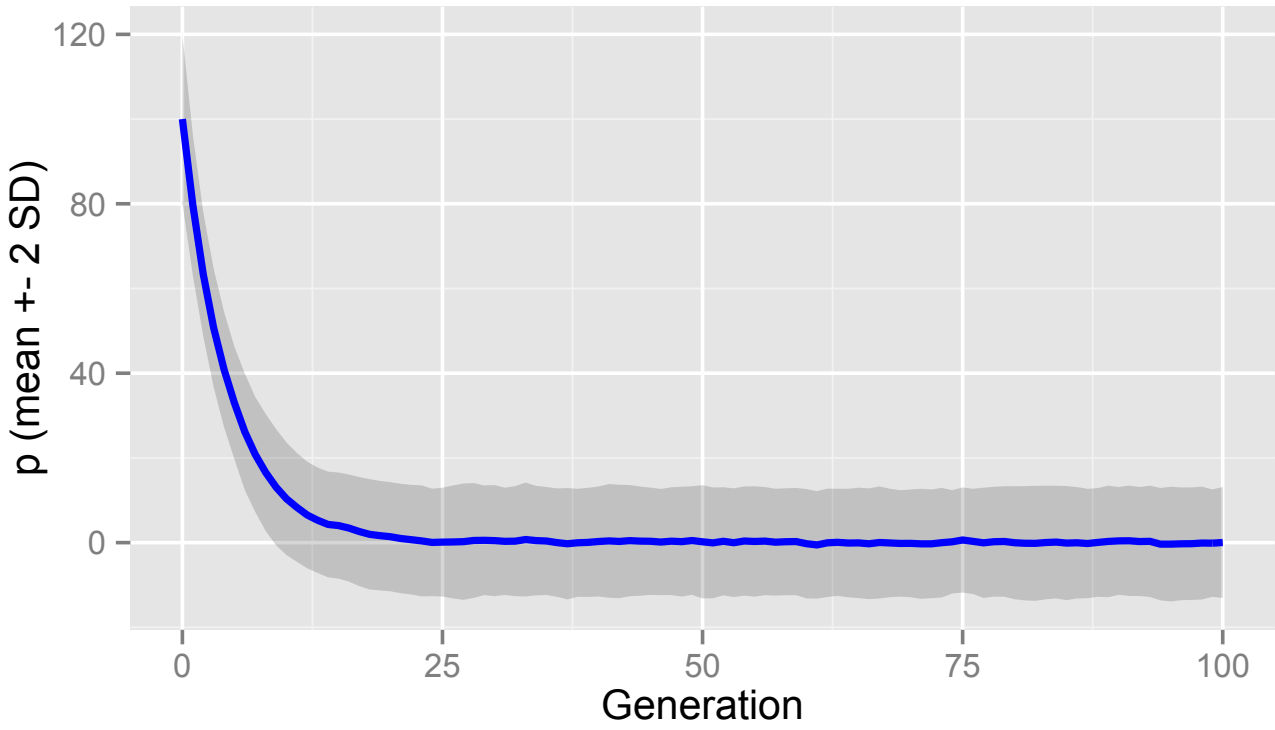
Framework

- *Population structure:* learners learn from (potentially) **multiple** teachers

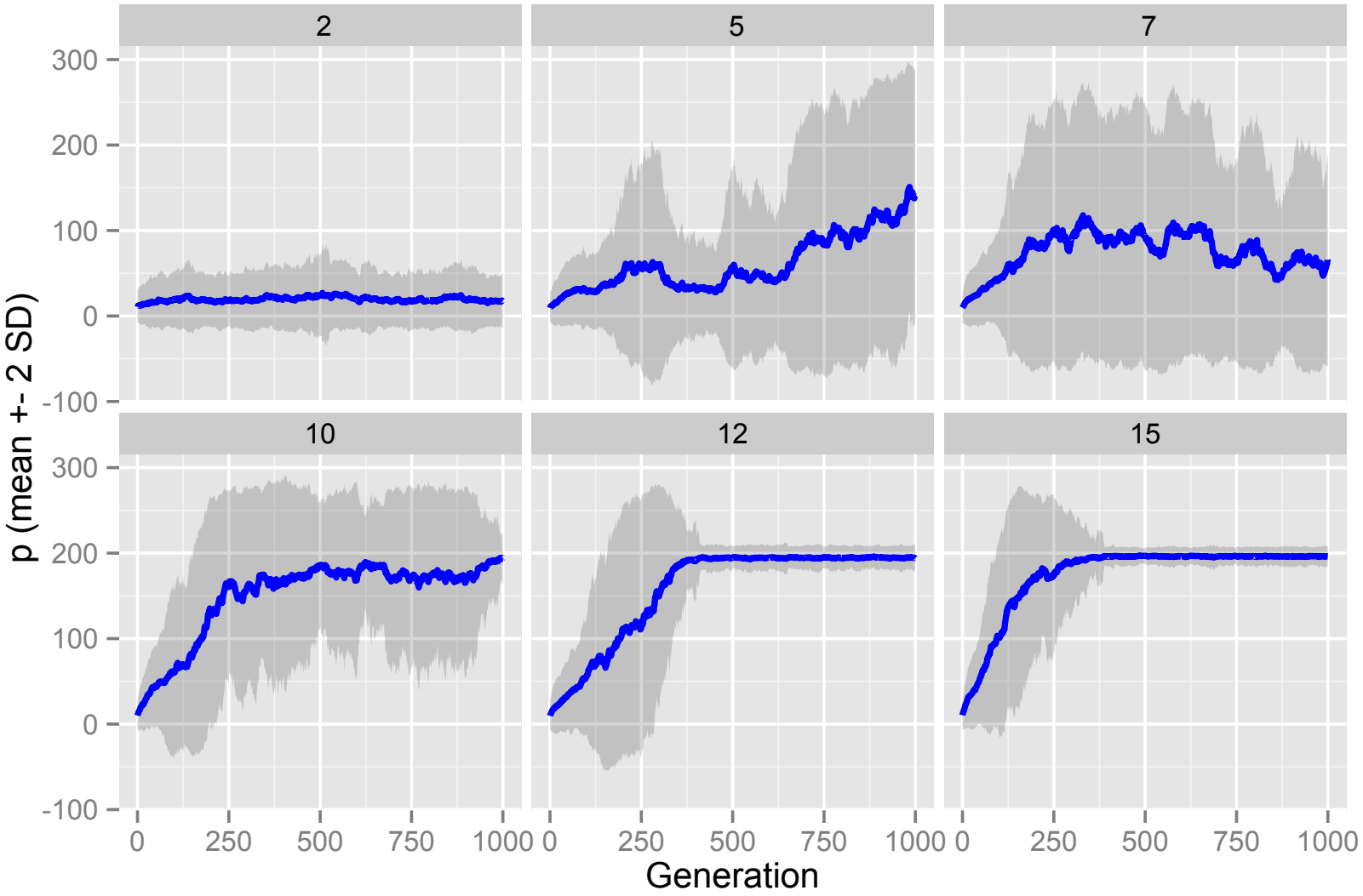


Framework

- *Outcome:* distribution of p in the population at time t ($\pi_t(p)$)



Effects of varying production bias (KS 2013, Model 3)



KS (2013)

- Only model with **both** production and categoricity biases could achieve all 3 goals:
  - stable limited coarticulation (low ℓ)
  - stable full coarticulation (high ℓ)
  - change from one to the other (medium ℓ)
- In models with categoricity bias, dynamics are not linear and phonologization is not inevitable (cf. Baker, 2008)

Now

- Production bias is the external force most commonly invoked in models of sound change
- ... but clearly not behind **all** changes: many other factors invoked by (socio)phon(eticians), e.g.
  - Contact (between subpopulations)
  - Social weight (of variants, speakers, groups)
  - Interaction (convergence, divergence)
- Today's questions:
  1. Does using production bias as the external force have a unique dynamics?
  2. If not, will *any* kind of external force produce the same behaviour at the population level?

Now

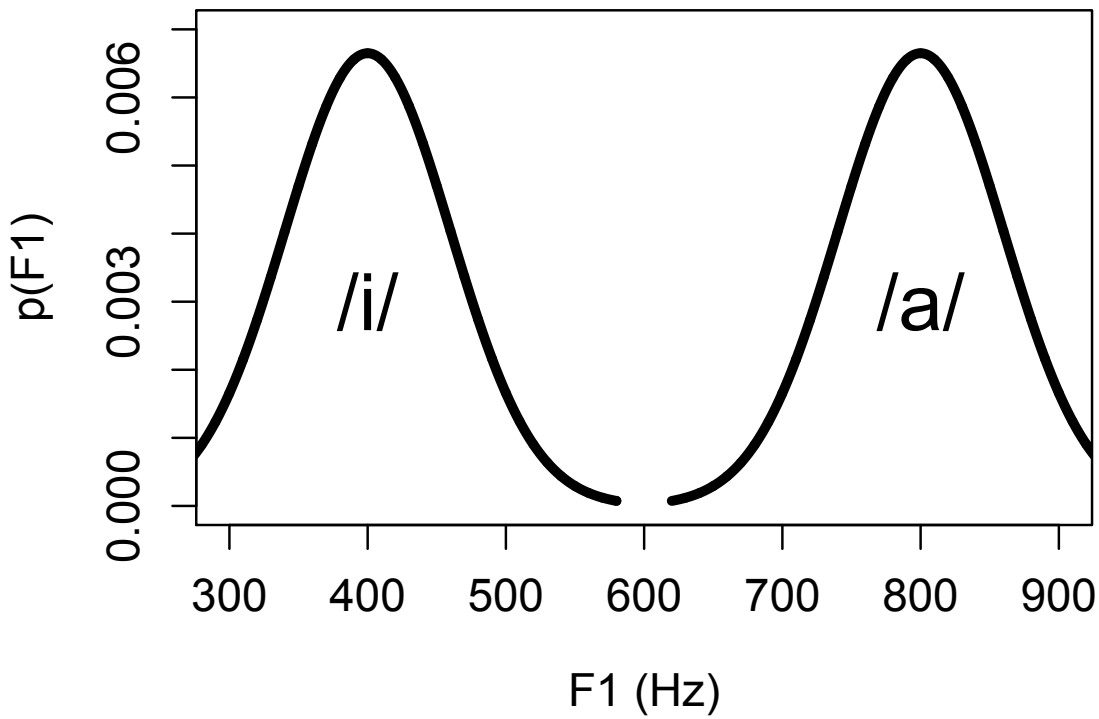
- Production bias is the external force most commonly invoked in models of sound change
- ... but clearly not behind **all** changes: many other factors invoked by (socio)phon(eticians), e.g.
  - Contact (between subpopulations)
  - Social weight (of variants, speakers, groups)
  - Interaction (convergence, divergence)
- Today's questions:
  1. Does using production bias as the external force have a unique dynamics?
  2. If not, will *any* kind of external force produce the same behaviour at the population level?

Subpopulations in contact: background

- Linguistic features can spread through contact between different groups (e.g. Thomason, 2001)
- These may be different languages, dialects, or subpopulations of a single group
- Are both stability and change possible when heterogenous groups interact?

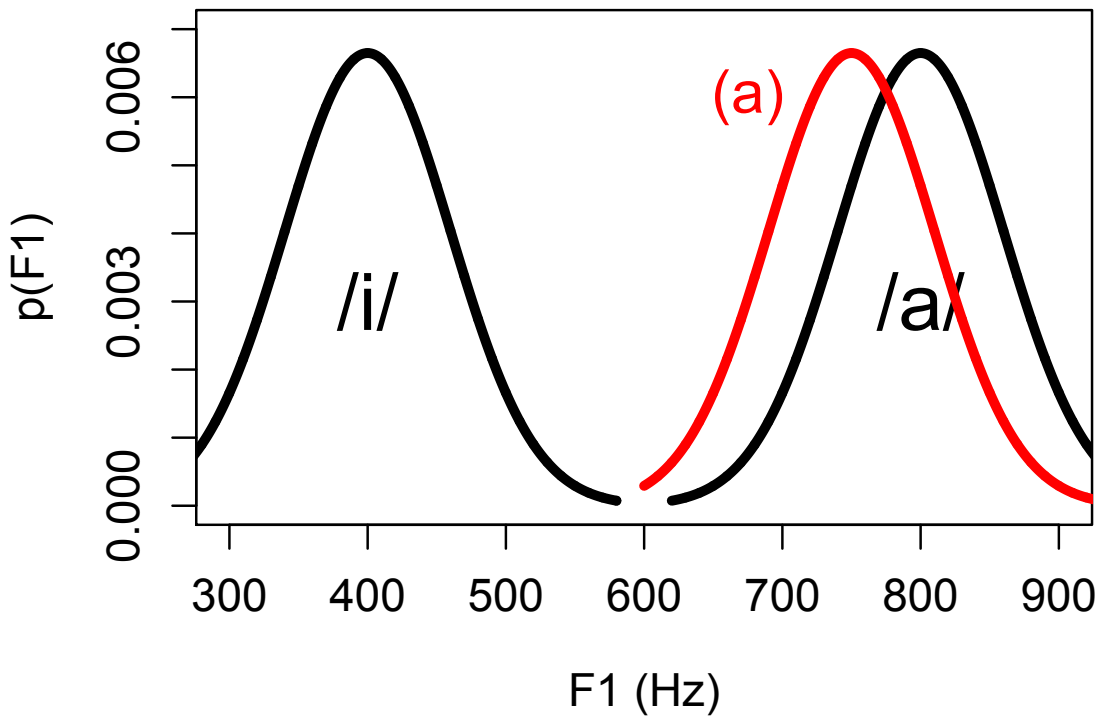
Model 1: Subpopulations in contact

- Simple instantiation: population divided into two groups:



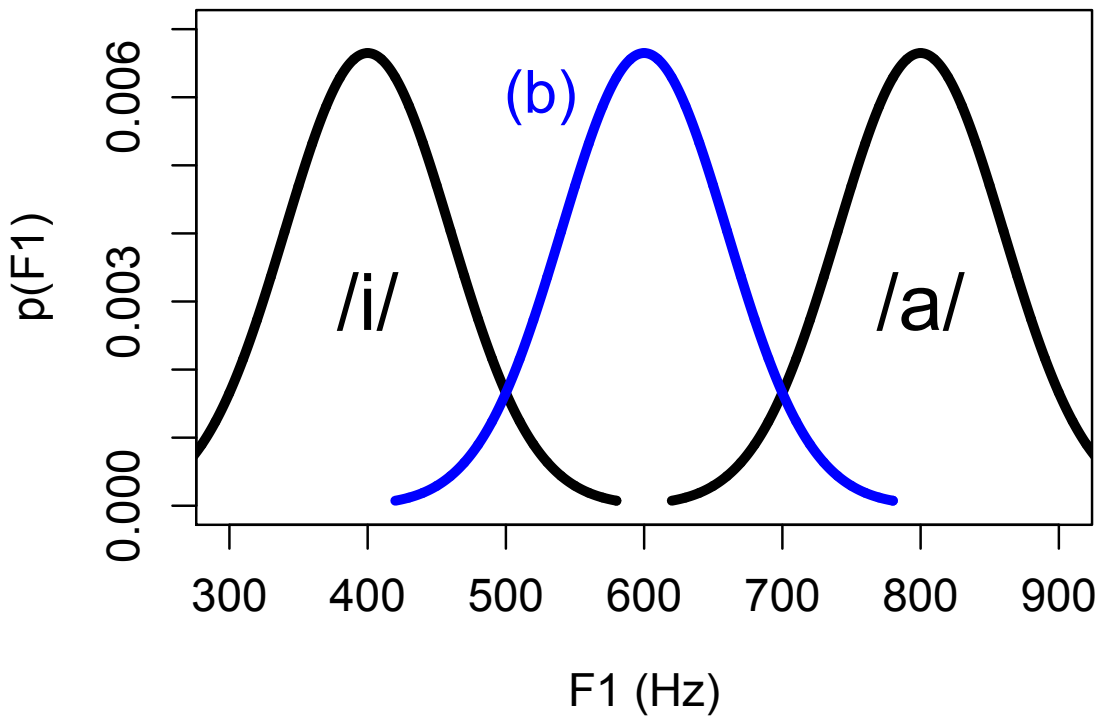
Model 1: Subpopulations in contact

- Simple instantiation: population divided into two groups:
  - Group *a* has little/no coarticulation



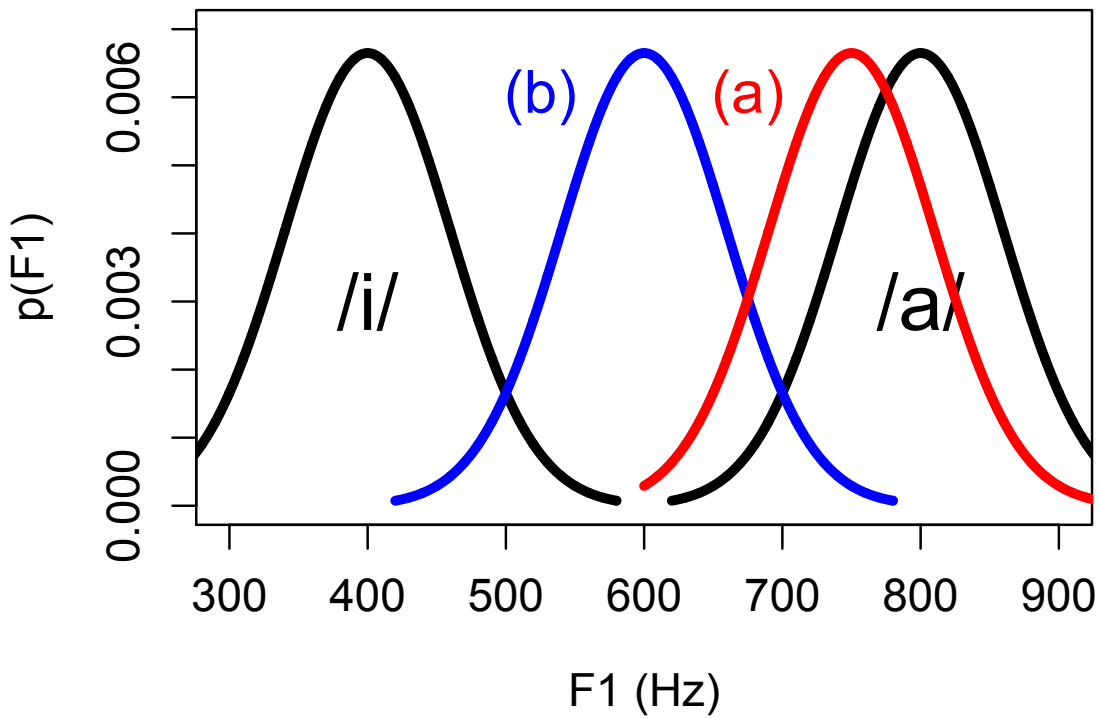
Model 1: Subpopulations in contact

- Simple instantiation: population divided into two groups:
  - Group *b* has extreme coarticulation

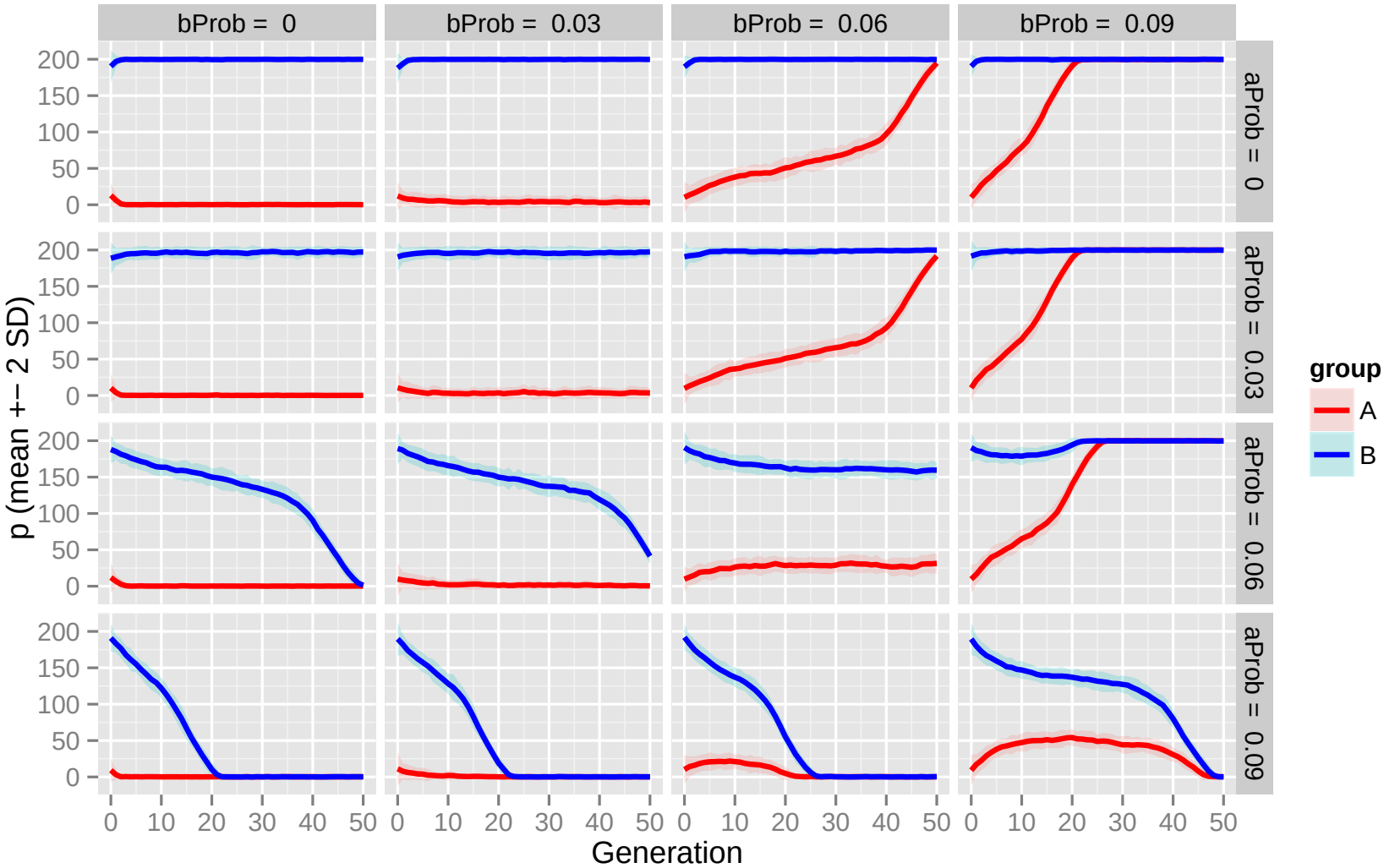


Model 1: Subpopulations in contact

- aProb: *P*(Group B agent learns from Group A agent)
- bProb: *P*(Group A agent learns from Group B agent)



Model 1: Results



Model 1: Discussion

- All three outcomes possible
- Stability can be preserved in both groups even when there is some interaction between them
- But: obtaining just 5% of training examples from a different group can be enough to induce the entire population to converge to a single group's mean

Social weighting: Background

- From the pool of synchronic variation, certain linguistic features can spread due to association with
  - particular variants
  - individuals
  - groups

  (e.g. Labov, 2001)

- Are both stability and change possible in the presence of social value associated with:
  - more coarticulated variants (nearer to [i])
  - speakers who coarticulate more
  - groups 〃 〃

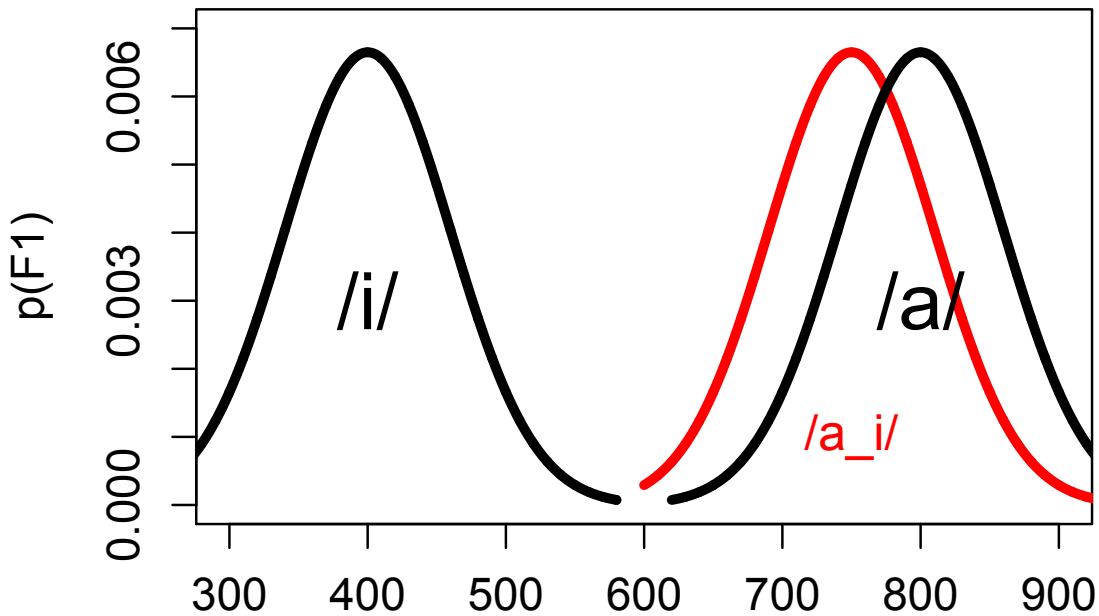
  ?

Models 2–4: social weighting

- Each token y_i has a social weight $w_i \in [1, w_{max}]$
- Higher social weight associated with:
  - Model 2: more coarticulated tokens (nearer to [i])
  - Model 3: tokens from teachers who coarticulate more
  - Model 4: tokens from high-coarticulation group
- Learner estimates p using weighted average of the y_i
  - tokens which are {more coarticulated, from teachers/group which coarticulate more} have more influence

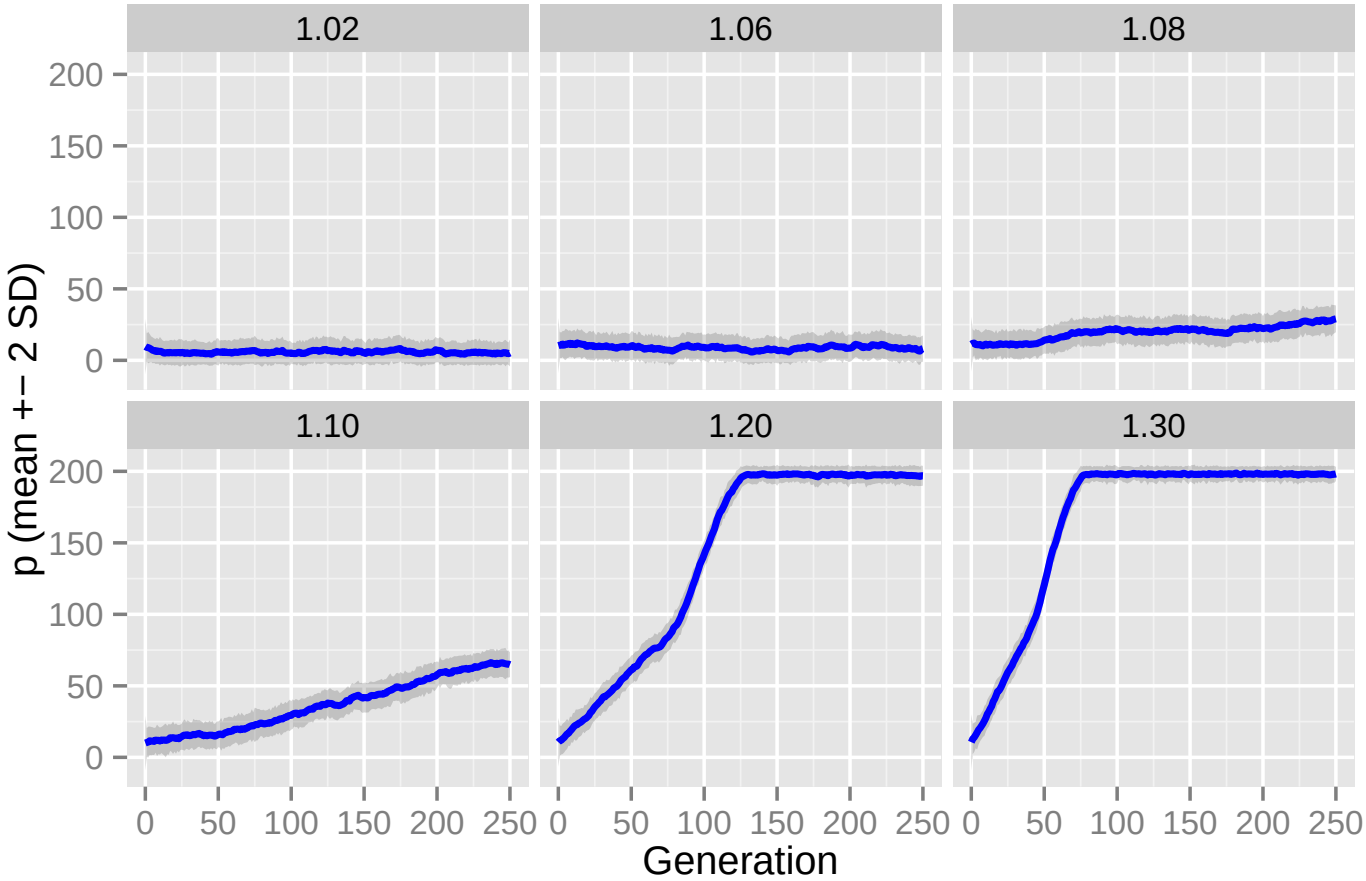
Model 2: social weighting by variant

- Start with a single population, little coarticulation
- Parameter: w_{max} (preference for coarticulated variants)



Model 2: Results

Varying w_{max}:

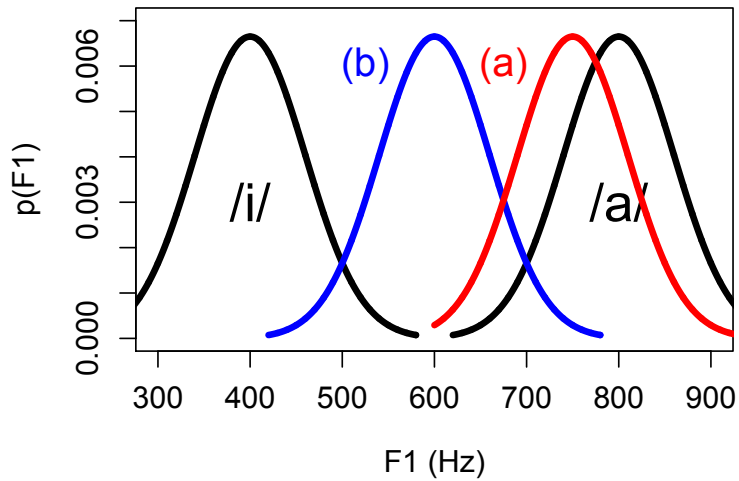


Model 2: Discussion

- All three outcomes possible
- Stability can be preserved even when coarticulated variants are socially valued
- But: social value of coarticulated variant just 10% more than uncoarticuated variant can be enough to induce change to full coarticulation in the whole population!

Model 3: social weighting by group

- Same architecture as Model 1:

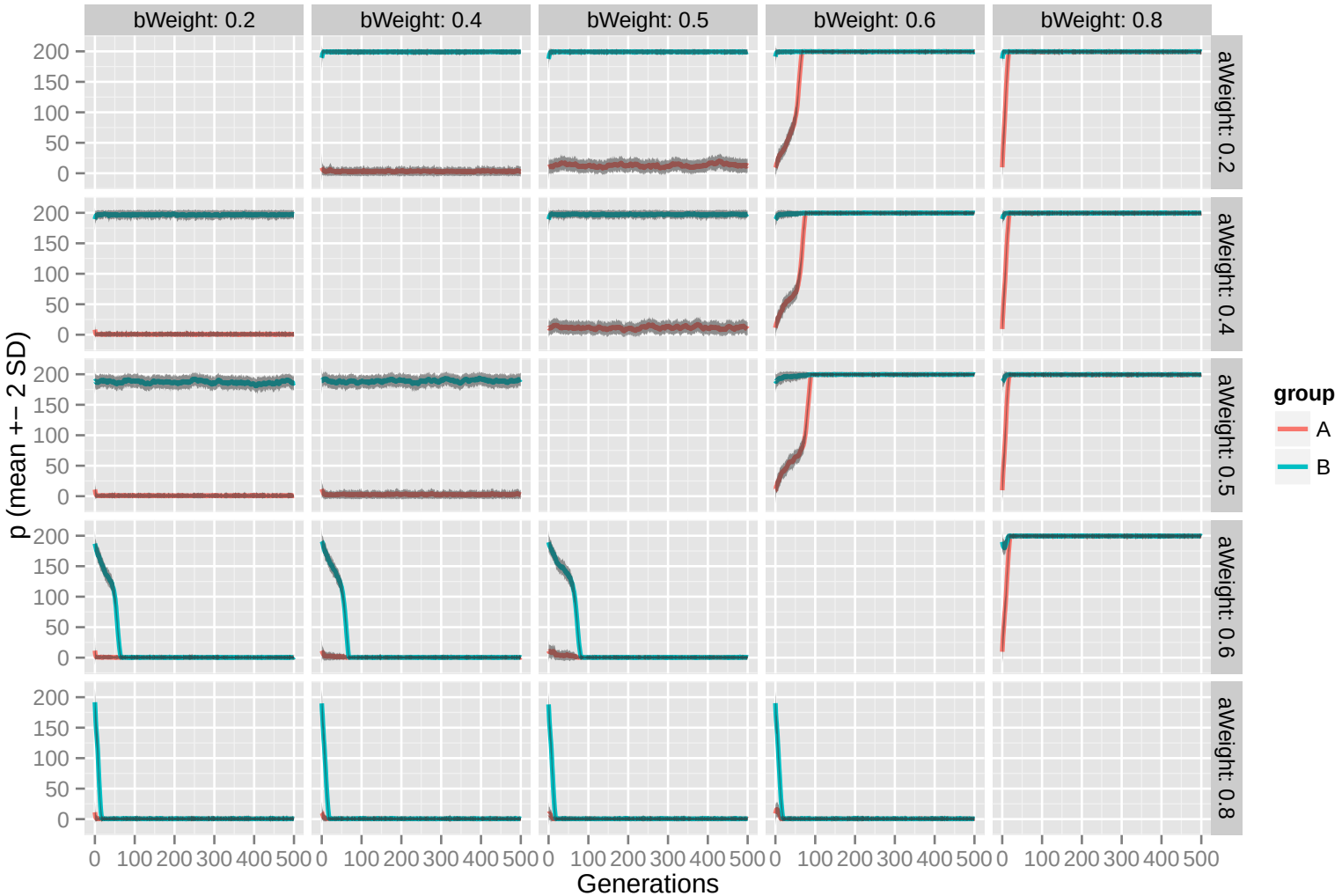


but with additional parameters: weight of

- data from group A for learner in group B: aWeight
- data from group B for learner in group A: bWeight

Model 3: Results

Fix aProb = bProb = 0.03



Model 3: Discussion

- All three outcomes possible
- Stability can be preserved even when group with high coarticulation socially valued
- But: even a small preference for tokens from coarticulating group can be enough to induce change to full coarticulation in the whole population.

Interim summary: Models 1–3

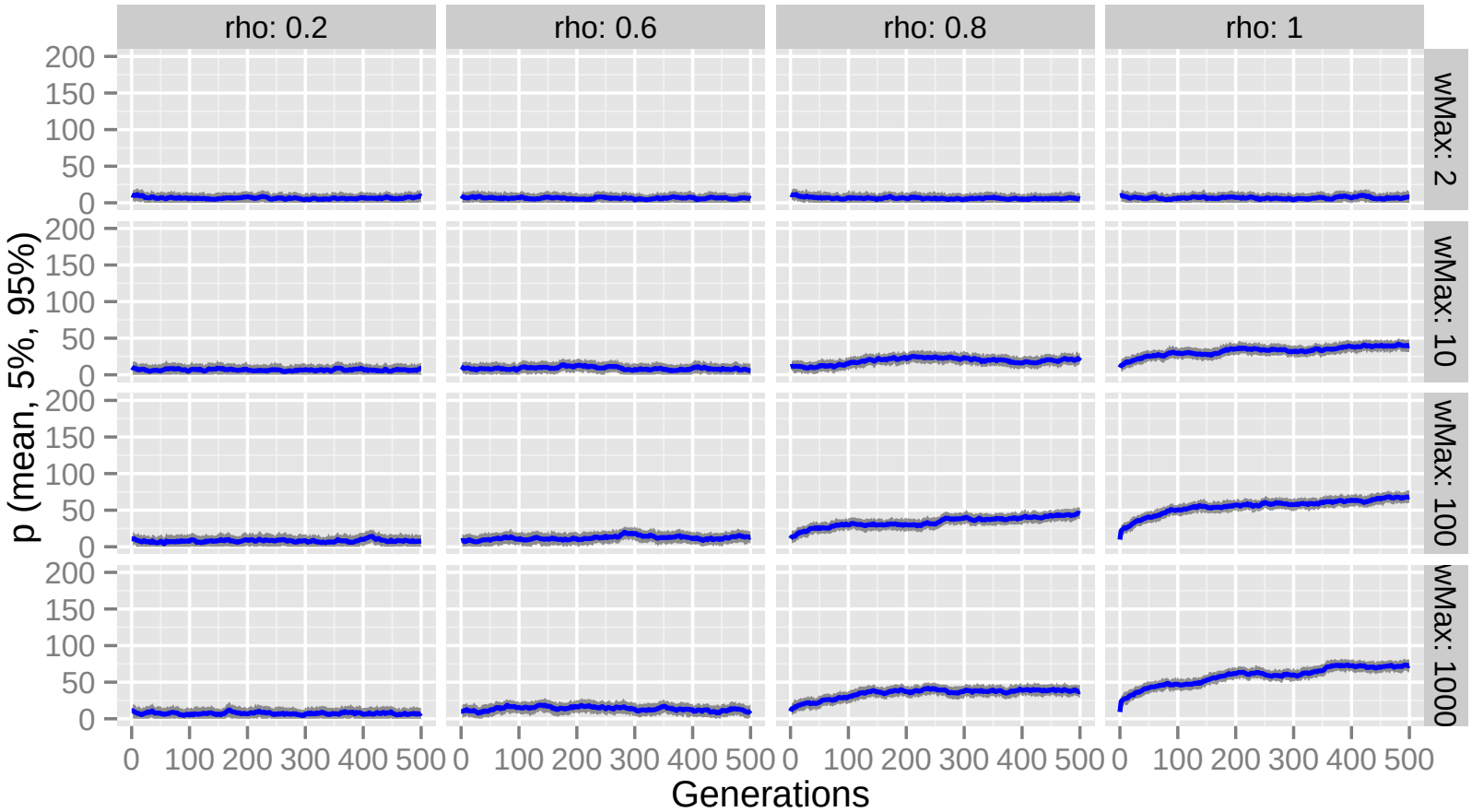
- Question 1: does driving force = production bias give unique dynamics?
- No: very similar dynamics when driving force is
  1. extent of contact
  2. social weighting of variants
  3. social weighting by group
- Question 2: will any kind of driving force produce the same behavior?

Model 4: social weighting by individual

- Setup: every teacher in generation t has
  - a social weight
  - a value of p
- If these happen to be positively correlated (i.e., data from teachers who coarticulate more is more highly valued):
  - more coarticulation in generation $t + 1$
  - could accumulate and lead to change

  (cf. Baker, Archangeli & Mielke 2011)
- Parameters:
  - w_{max}: maximum social weight
  - ρ: correlation between teacher's prestige and degree of coarticulation

Model 4: Results



Model 4: Discussion

- Stability: default
- Change: not really
- Driving force is much weaker than Models 1–3! “Change”:
  1. requires near-perfect coarticulation/social weight correlation, individuals who coarticulate weighted 100-1000x higher than individuals who don’t.
  2. is very slow (1000s of generations)
- Compare: change in $<$ 200 generations for small increases in driving force in Models 1–3

Model 4: Discussion

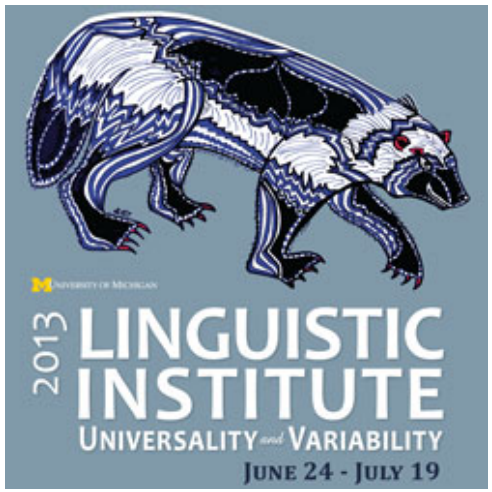
- Models 2– 4: all implementations of social weight. Why are dynamics of Model 4 different?
- Social weight on individuals (M4):
  - correlation between w and observations: weak
- Social weight on groups (M3):
  - correlation between w and observations: stronger
- Social weight on variants (M2):
  - correlation between w and observations: perfect
- Question 2: will any kind of driving force produce the same behavior?
  - No

Conclusions

- different external forces + categoricity bias = similar population dynamics
  - Implication: a similar dynamics may underlie actuation of changes initiated from different sources
  - Good: sound change can have different sources, and doesn't show radically different dynamics by source (?)
- But not all external forces give both stability and change
  - Some intuitively plausible mechanisms "too noisy" to have an effect iterated over time in a speech community.
  - Population dynamics as partial solution to the "non-phonologization problem"

Thanks!

- Ideas/comments:
  - Participants in “Computational Models of Sound Change” at 2013 LSA Institute
  - Audiences at Ohio State, McGill
- Funding: FRQSC #183356, CFI #32451



Fonds de recherche
Société et culture
Québec

INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION