Evolution: Comparing Biology and Culture

Andy Wedel
Department of Linguistics
University of Arizona

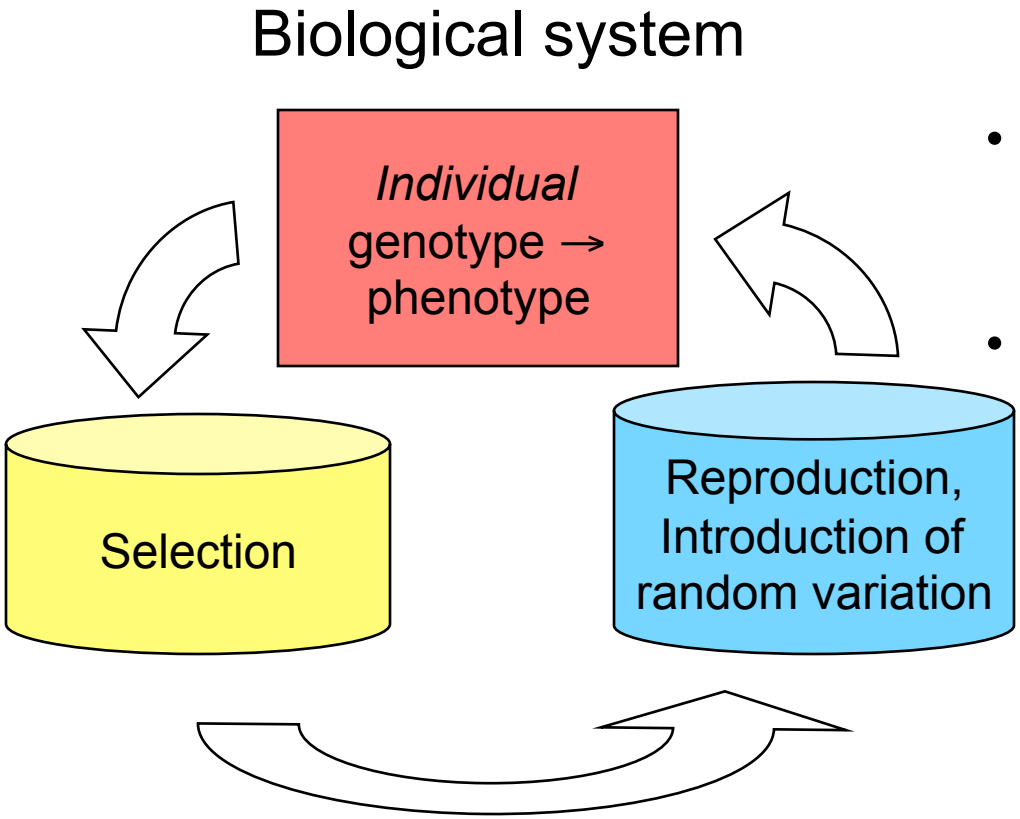
Evolution as a population-based mechanism of change

- Requirements for evolutionary change in the frequency of a trait:
  - *Population* of *replicating elements*
  - Trait influences relative probability of replication
  - Trait is *heritable* to some degree

Some questions relevant to evolution of mental representations (Henrich, Boyd, Richerson 2002)

1. Do sources of variation need to be random?
   - Does ‘selection’ need to be separate from the generation of variation?
2. Do replicating entities need to be discrete?
   - How important is the literal idea of a *population*?

Selection in biological systems



- Selection is functionally decoupled from production of variation.
- No Look-ahead
  - Interesting limitation on ability of system to explore possibility space.

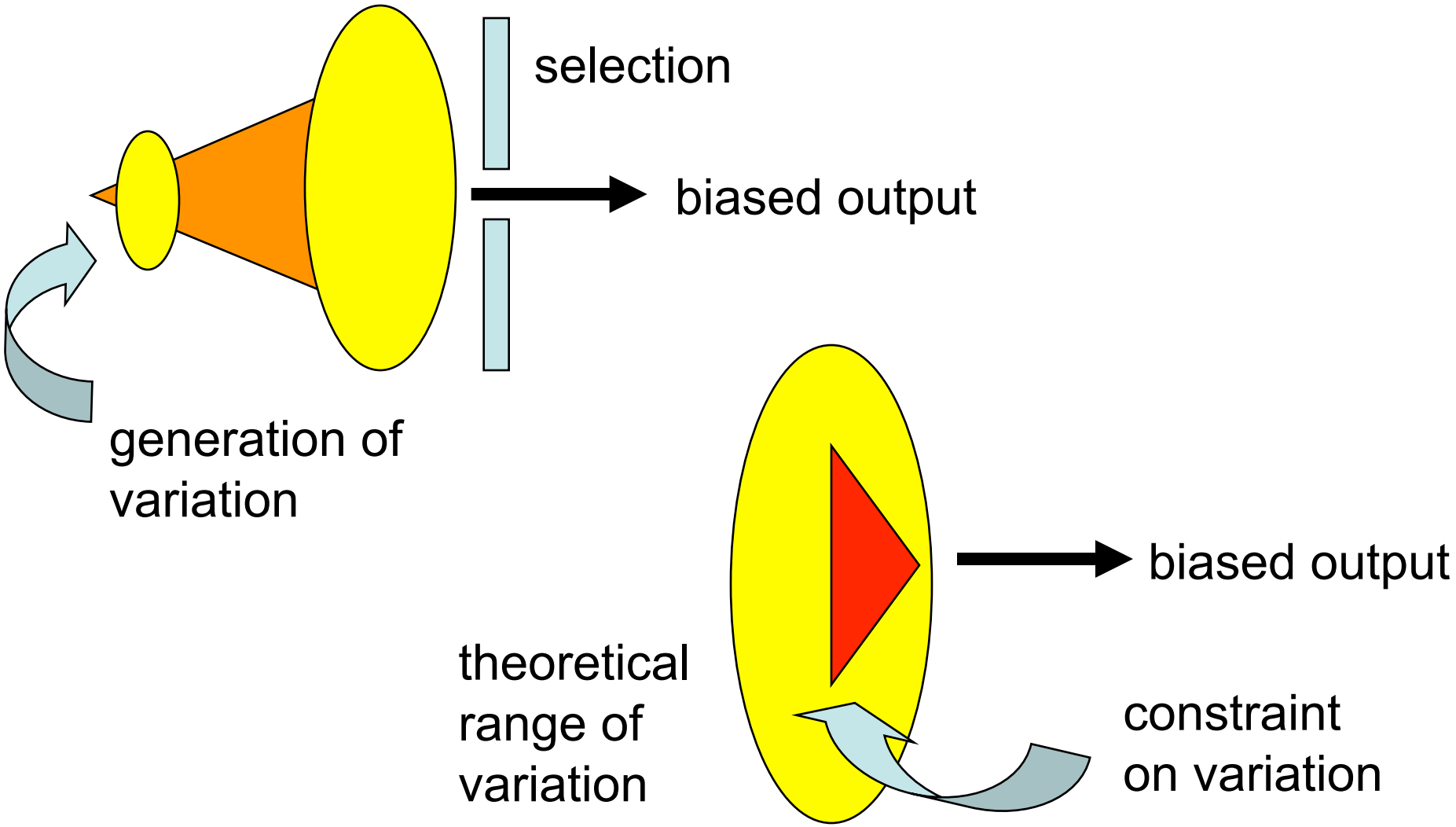
Must variation be random?

- What's the situation in biological evolution?
  - Variation is highly constrained
  - But random with regard to phenotype
- In linguistic evolution?

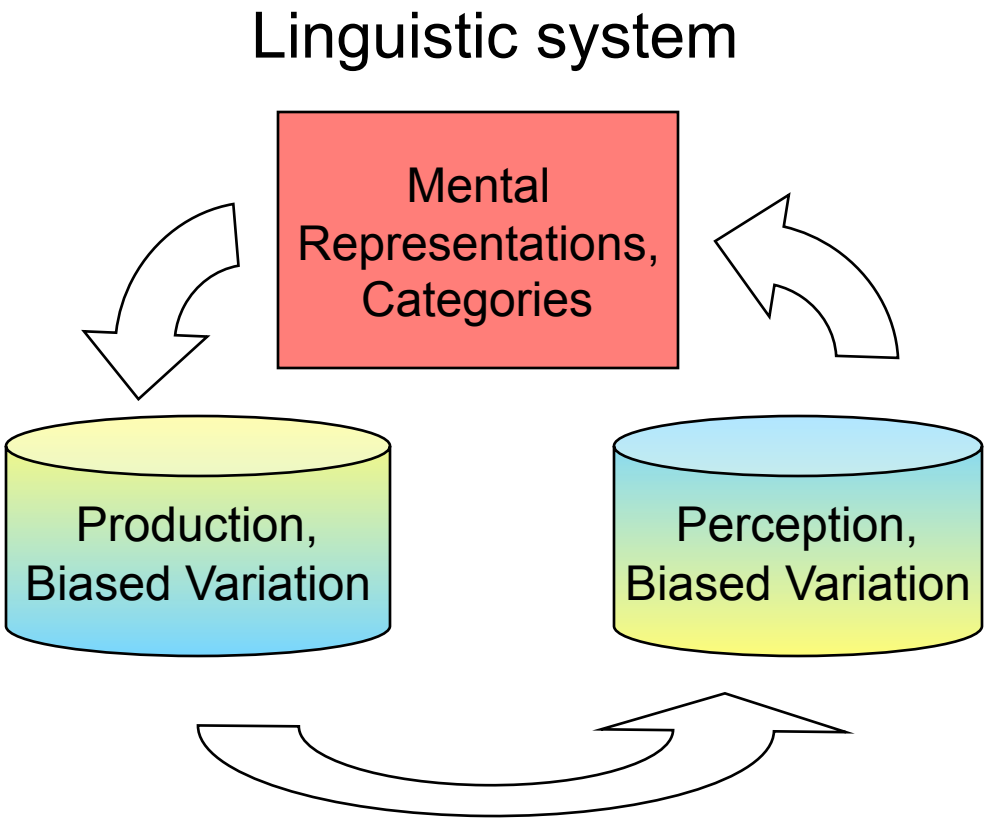
Variation in language: *not* random with regard to 'phenotype'

- Production
  - Articulation, aerodynamics
- Perception
  - salience, acoustic intensity
- Categorization
  - Structure preservation
    - 'thorn' example
    - Austronesian example (Blevins, in press)
  - Creates feedback loops; biological analogy in sexual selection

Thought experiment



Biased variation as selection in cultural evolution



- Selection can operate *through* the biased production of variation.
- Provides a limited ‘look-ahead’:
  - The current state of the system can influence error in production and perception

Example: Model of contrast evolution in sublexical categories.

- No discrete replicators
  - Every member of the model population contributes to some degree to every output
- Selection arises through biases in variation, not through biases in survival.

Contrast in Sublexical Categories

- To the extent that words are composed of smaller units, in order for words to be contrastive, the set of smaller units must themselves be contrastive.
  - Languages do have sets of contrastive sound categories, e.g., phoneme inventories.

Questions

- How is sound category contrast maintained through the course of sound change?
  - We know phonemes can be lost or merged.
  - But we also know that sounds often seem to change as if contrast were important.
    - Contrast trading
    - Chain shifts
    - Contrast maintenance (homophony avoidance) in paradigms
- Range of hypotheses:
  - intervention by innate monitor of contrast
  - epiphenomenon of language change
  - indirect result of contrast function

Homophony avoidance in Trigrad Bulgarian (Stojkov 1963)

a.	<u>sg. /-o/</u>	<u>pl. /-a/</u>	
	zórn-o	zórn-a	‘grain, seed’
	pétal-o	pétal-a	‘horseshoe’
	blág-o	blág-a	'blessing'
	cigaríl-o	cigaríl-a	'cigarette'
b.	kapít-a	kapit-á	‘hoe’
	kláb-a	klab-á	‘ball of thread’
	pér-a	per-á	‘feather’
	rébr-a	rebr-á	‘rib’

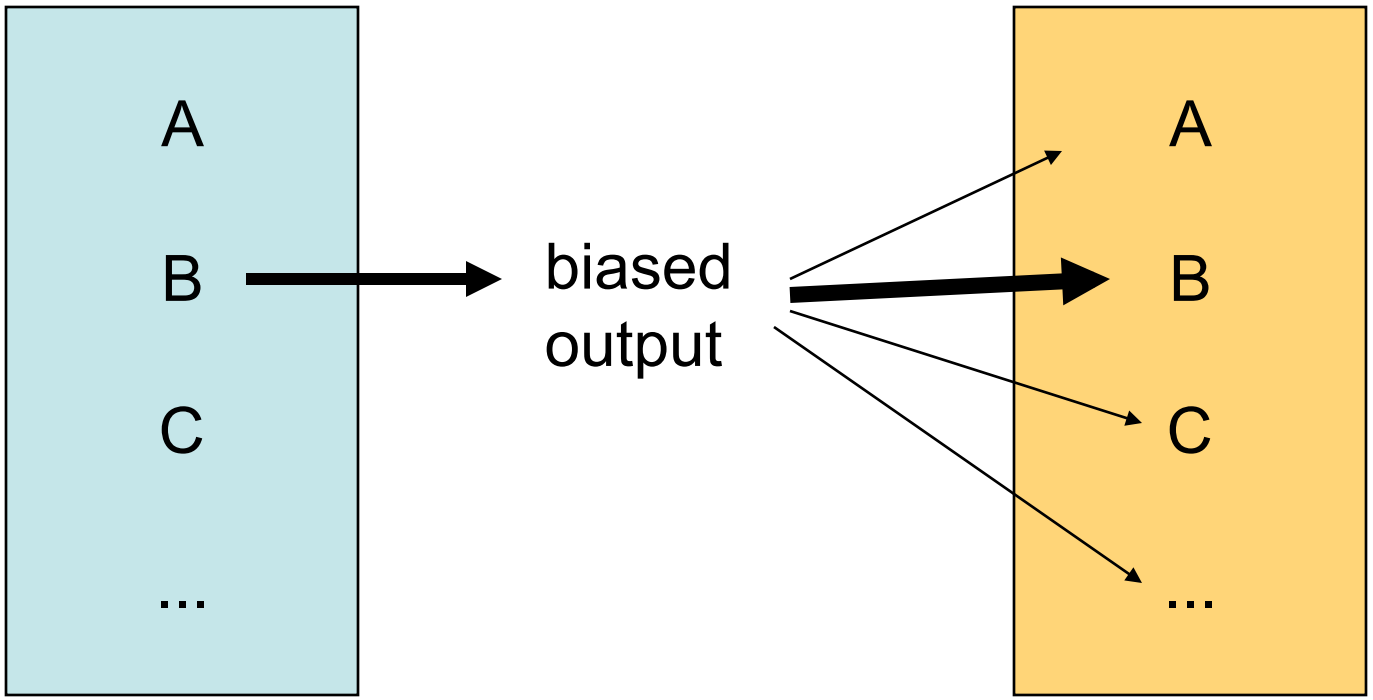
Build a model

1. Illustrates feedback from selection for lexical contrast to promote system of sound contrasts
2. Illustrates evolutionary system in which
   - There are no discrete replicators.
   - Selection is at the level of biased variation, not biased survival.

Model architecture

- Two (or more) agents
- Each has a fixed lexicon
  - Lexical entries contain exemplars of previously perceived words.
- Word exemplars consist of ordered sound exemplars.
- Two 1-dimensional sound continua (0-100)
  - Think VOT, or vowel height
  - Words are built from alternating values on these continua: CVCV...
  - example: **20 58 23 62**

Model Architecture

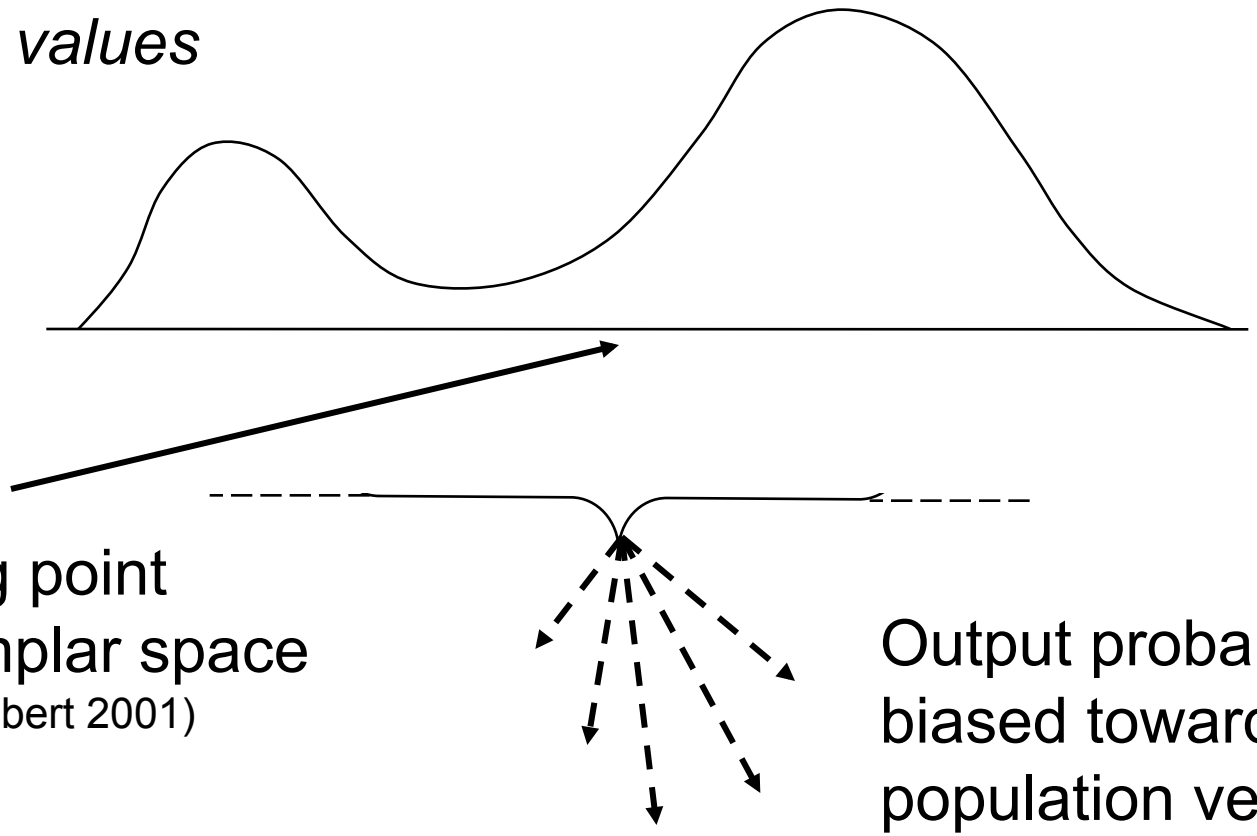


Two bias-types are included

1. Lenition: A random Gaussian biased toward the center point of the continuum (50) is added to sound values of outputs.
2. Output sound values are biased toward local peaks in the stored sound distribution. (Guenther and Gjaja 1996, Oudeyer 2006).

Production bias toward previously perceived sound values results in *reversion to the mean*.

Record of previous sound values

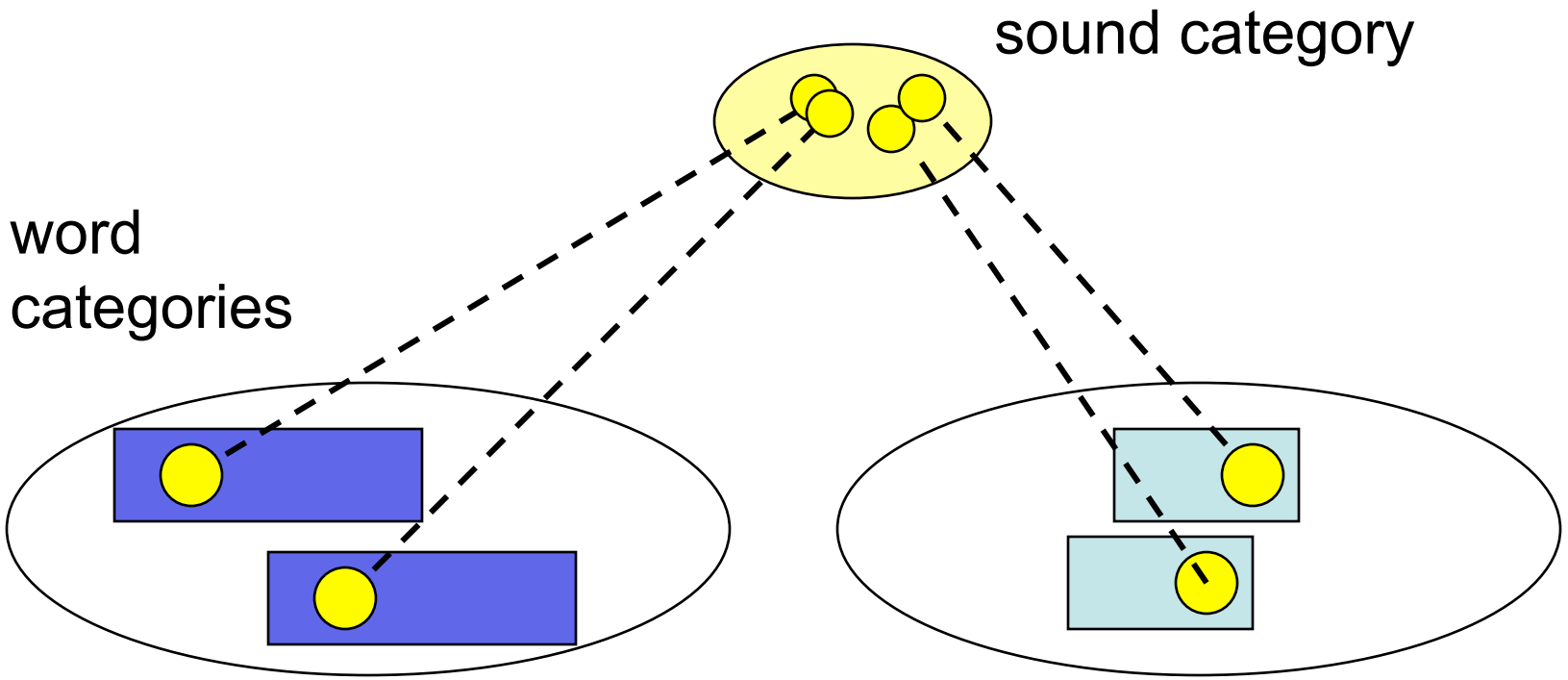


Starting point in exemplar space (Pierrehumbert 2001)

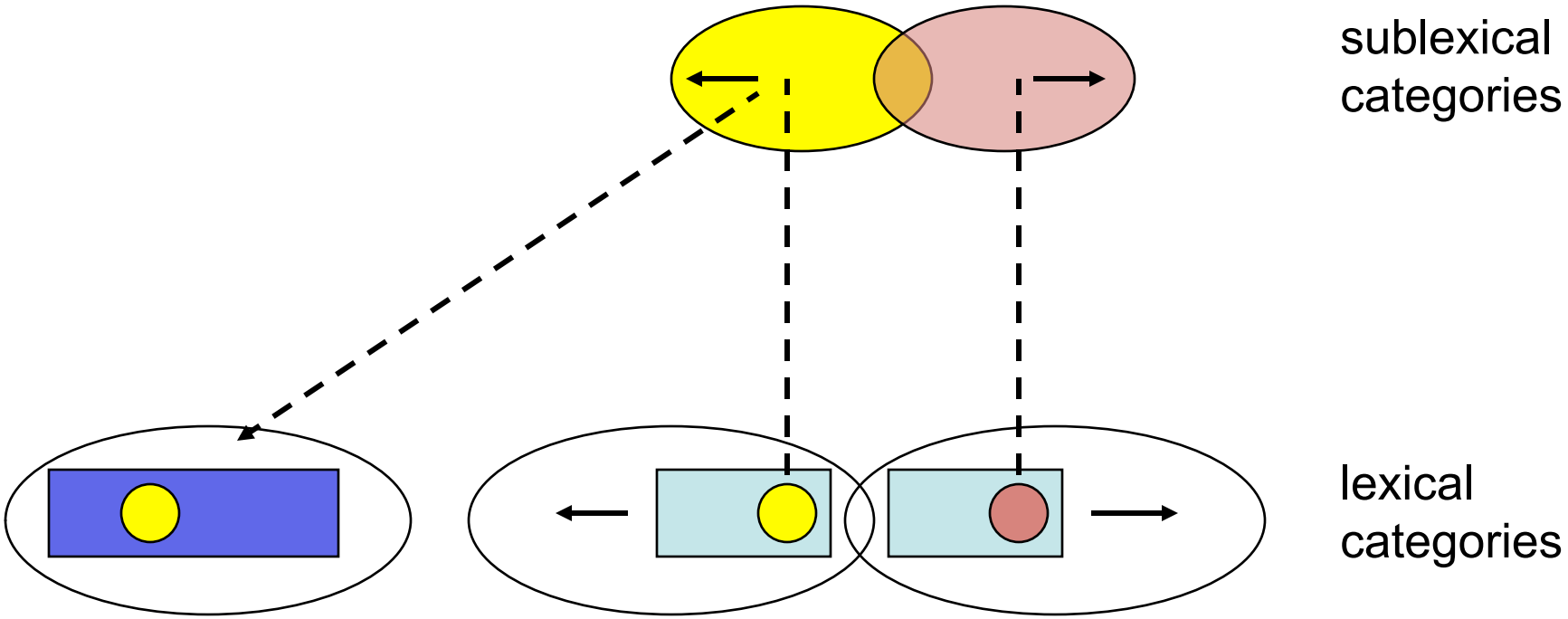
Output probabilistically biased toward the global population vector from that point (Oudeyer 2006)

Two interacting levels of categorization

- Words are composed of an ordered set of sounds.
- Sound and word categories consist of cross-referenced exemplars (e.g., Bybee 2001).
- Production involves blending at both categorial levels.



Change in words influences change in sounds



Conceptually parallel to individual:gene relationship

- Individuals contain genes.
- Selection is at the level of the individual
  - The entire set of an individual's genes are transmitted, or not.
  - Fitness is context-dependent
- Gene variants can spread through the population *even if they are only selected for in a subset of contexts*.

Two initial controls

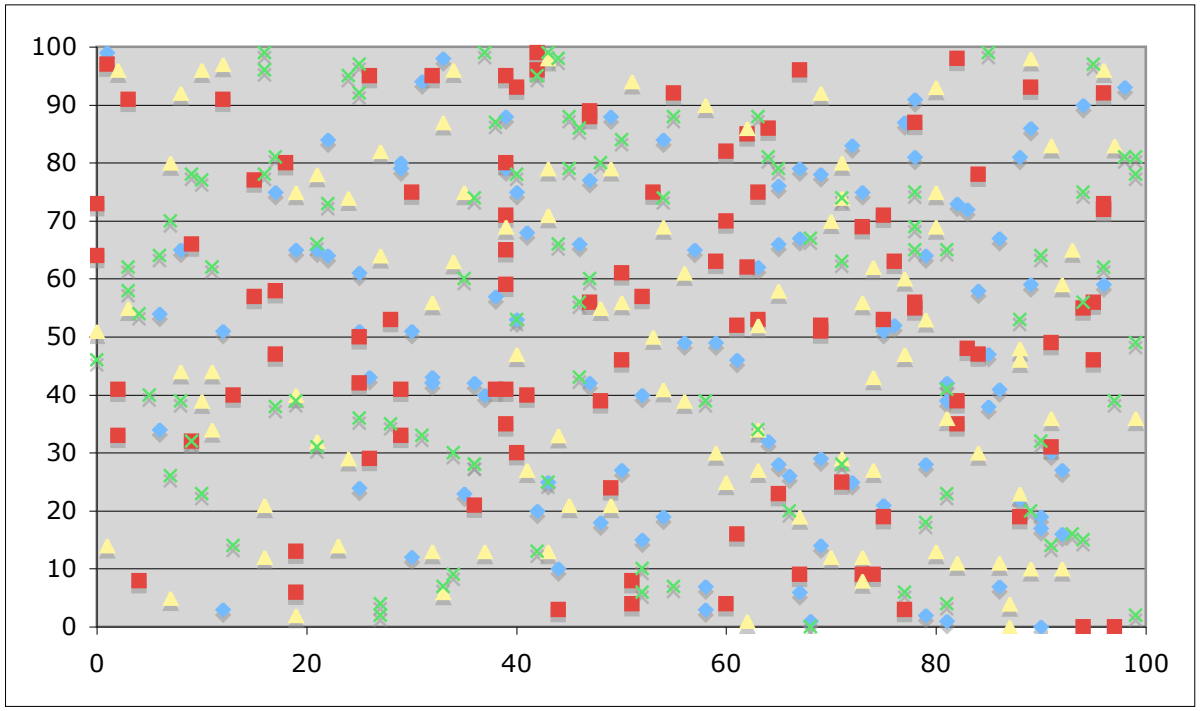
1. No competition between categories in the hearer
   - Removes selection for contrast at lexical level. How do sound distributions evolve without this selection for lexical contrast?
2. No reversion to the mean at the sound level.
   - Every word category evolves independently.

1. No selection for lexical contrast

- 4 CV word categories
- Begin simulation with randomly seeded lexical exemplars
- Run 4000 rounds, storing each output in the category intended by the speaker.

Cycle 0

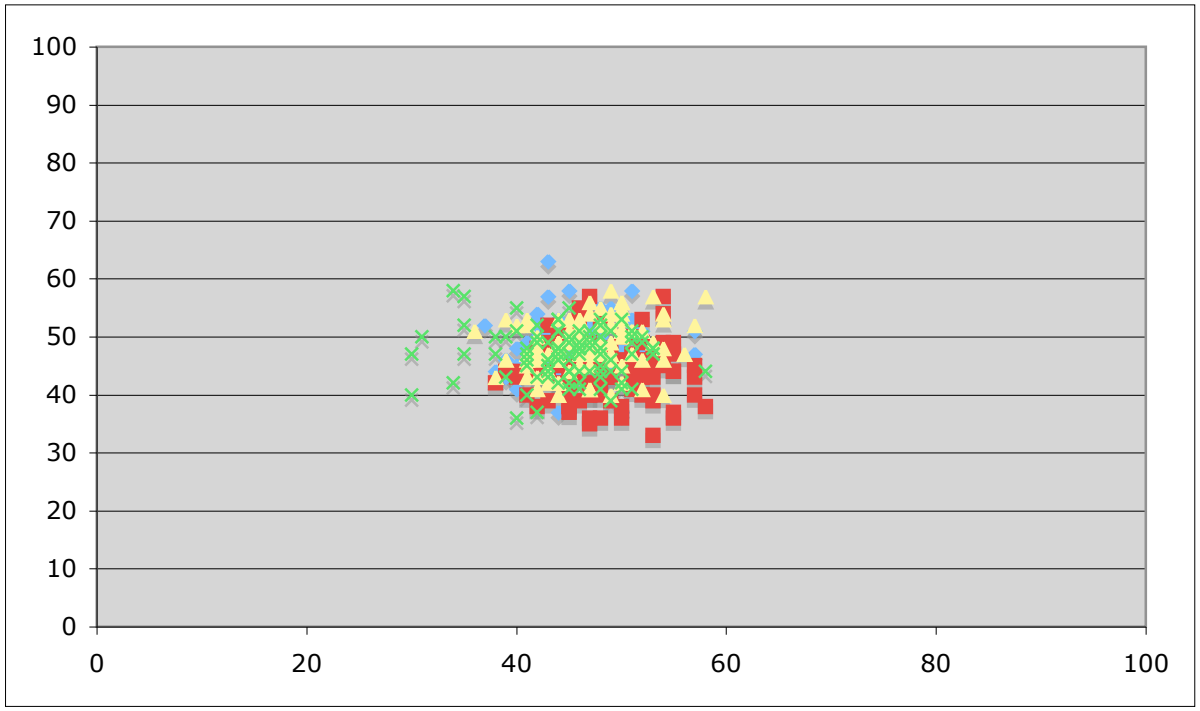
C



V

Cycle 4000

C



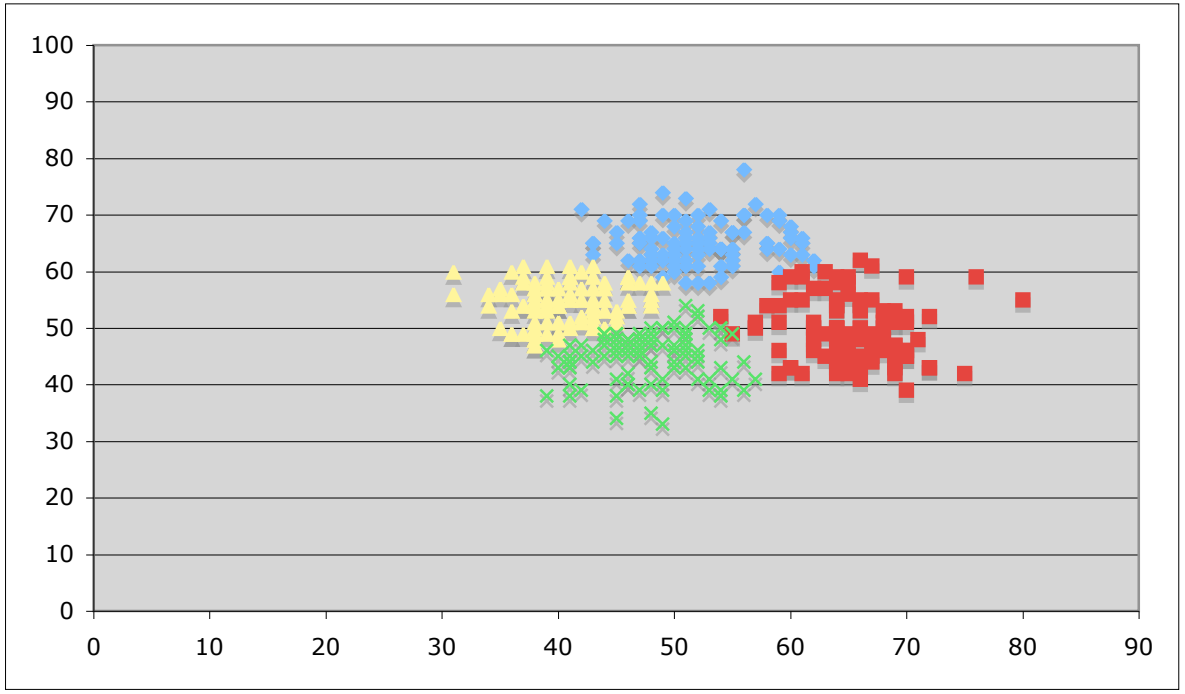
V

2. No reversion to the mean within sound distributions

- 4 CV word categories
- Begin simulation with all categories in the center of the C and V sound distributions.
- Run 4000 cycles

Cycle 1000

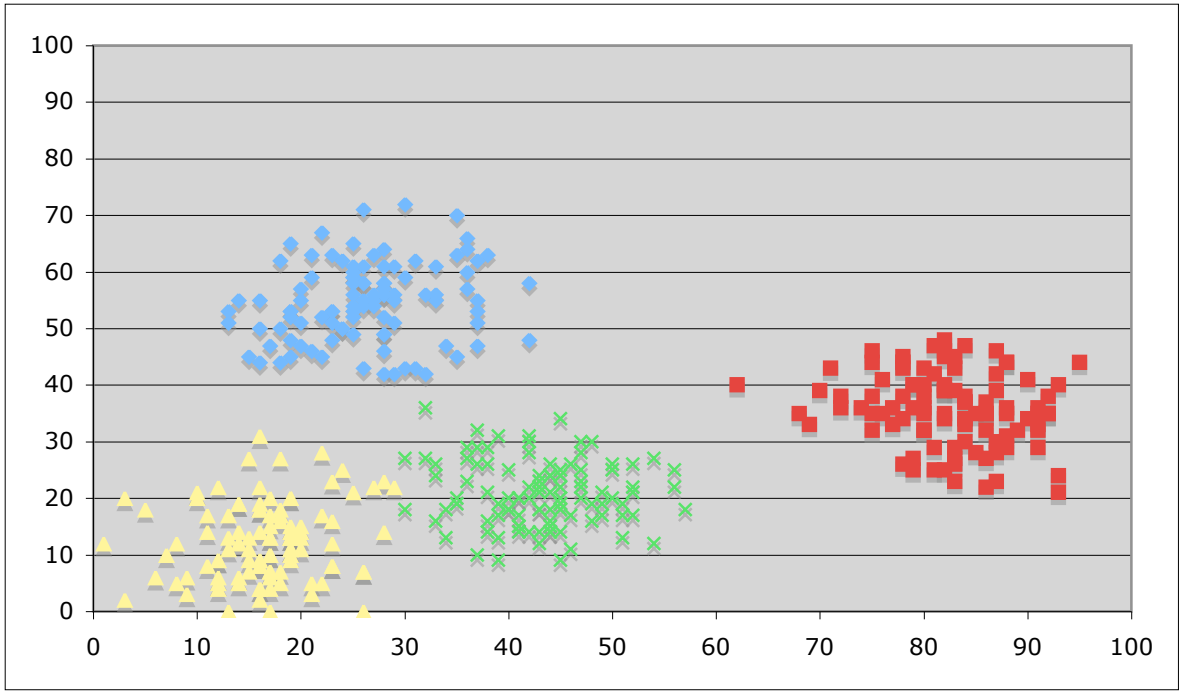
C



V

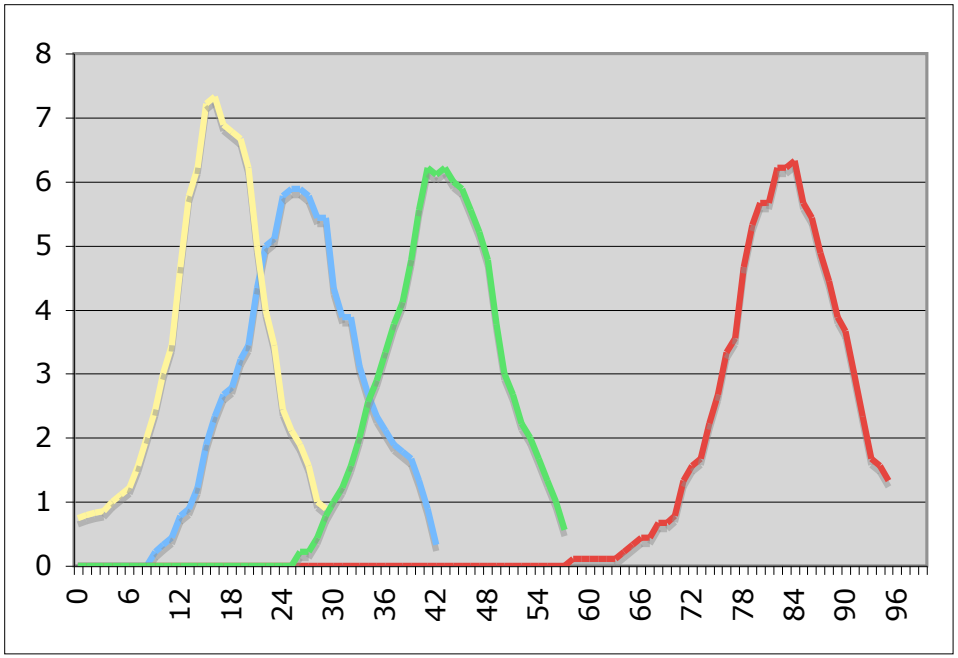
Cycle 4000

C



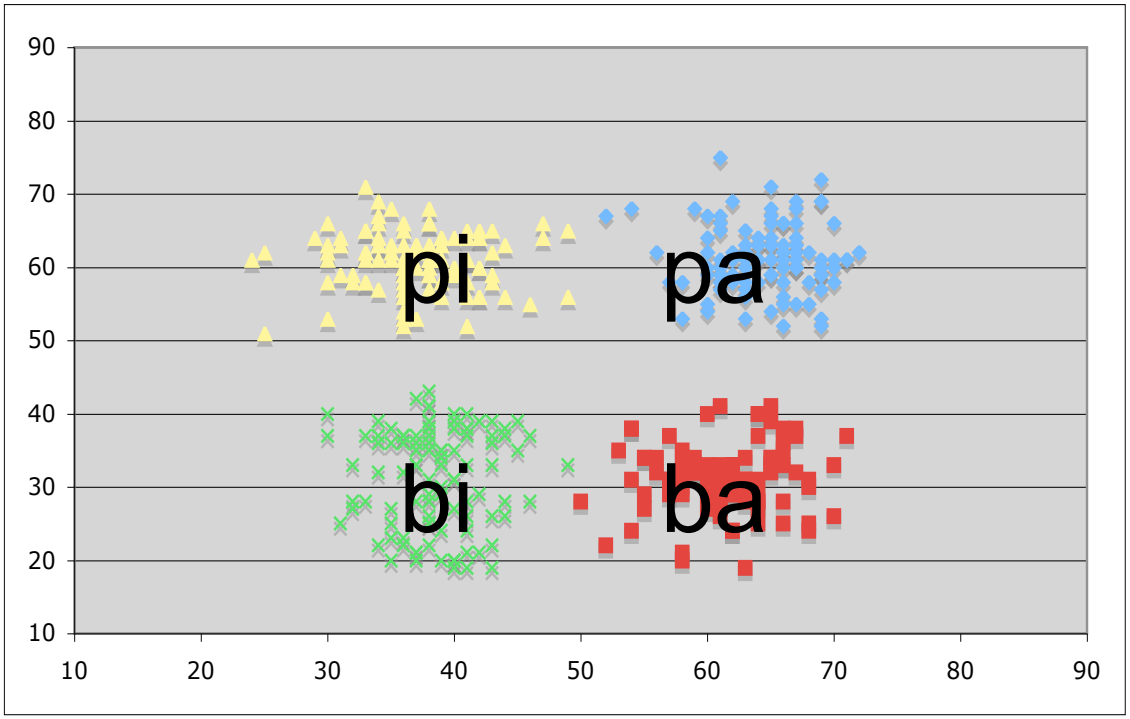
V

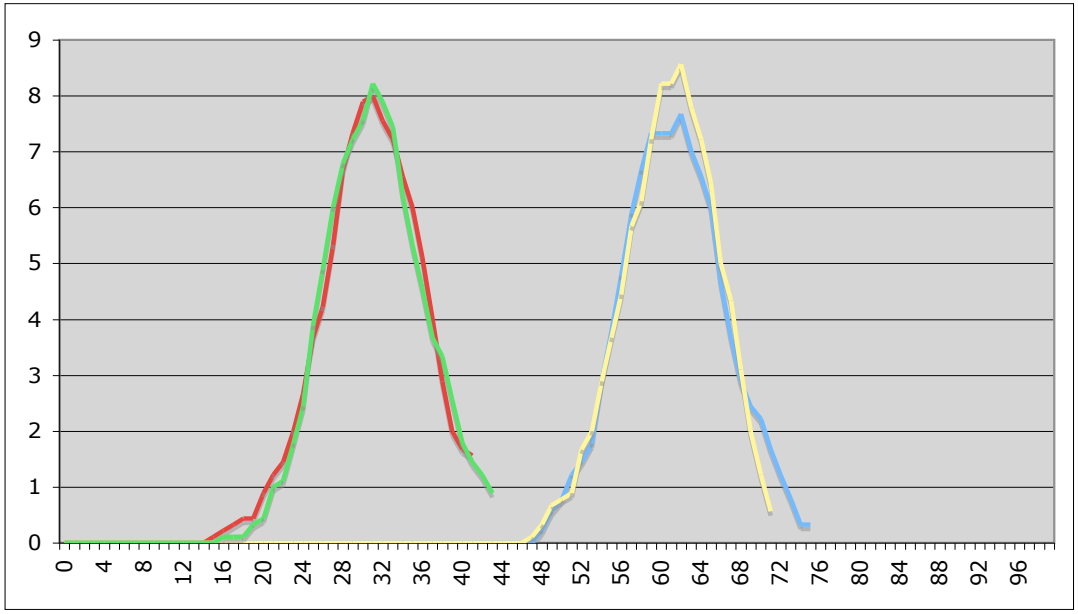
Distribution of V sound exemplars from each lexical category



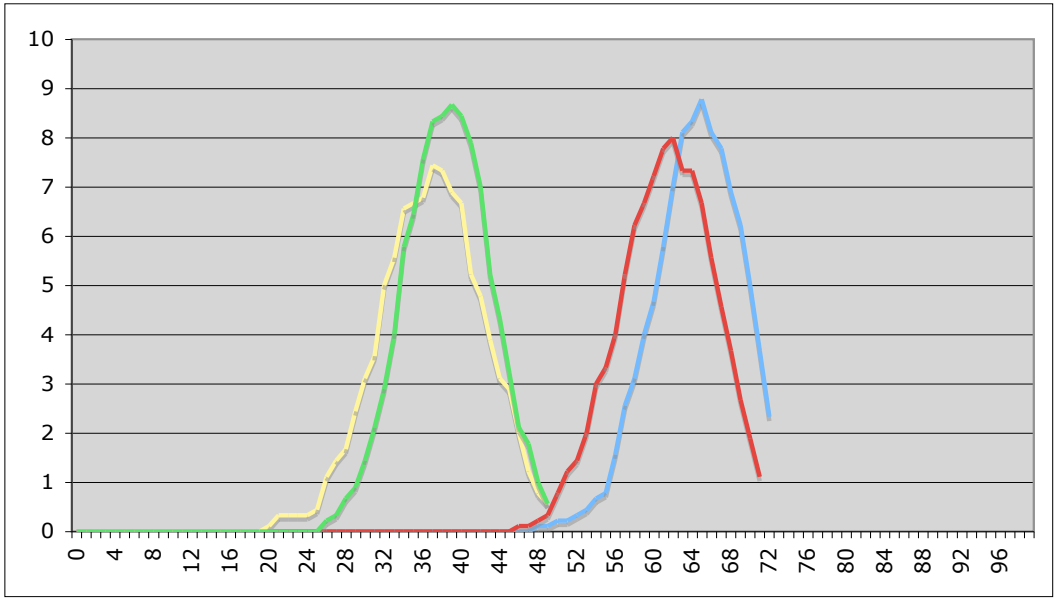
V

Add back reversion to the mean within sound distributions




9
8
7
6
5
4
3
2
1
0
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96

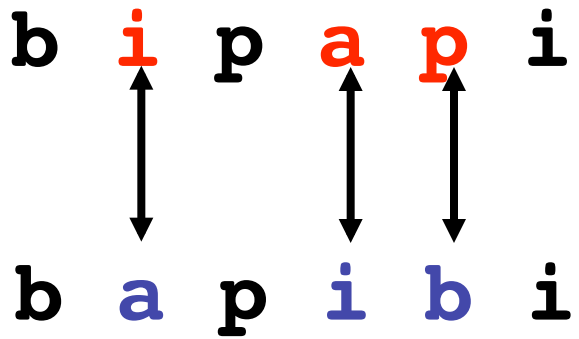
C


10
9
8
7
6
5
4
3
2
1
0
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96

V

10 CVCVCV lexical categories

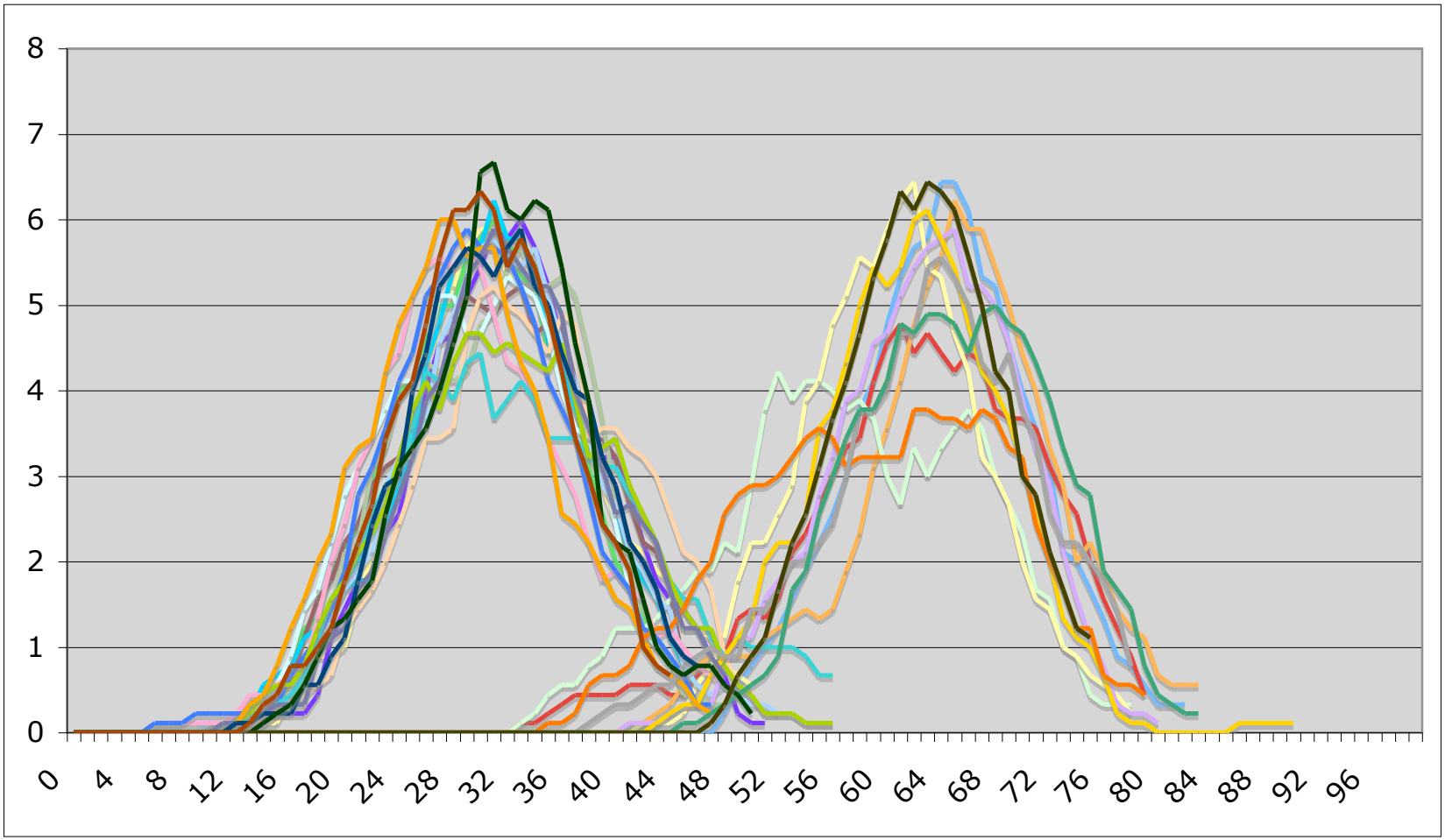
- 2 C and 2 V categories more than sufficient for contrast between all items.
  - For most lexical items, inter-lexical contrast is provided at *multiple* positions
  - e.g., compare



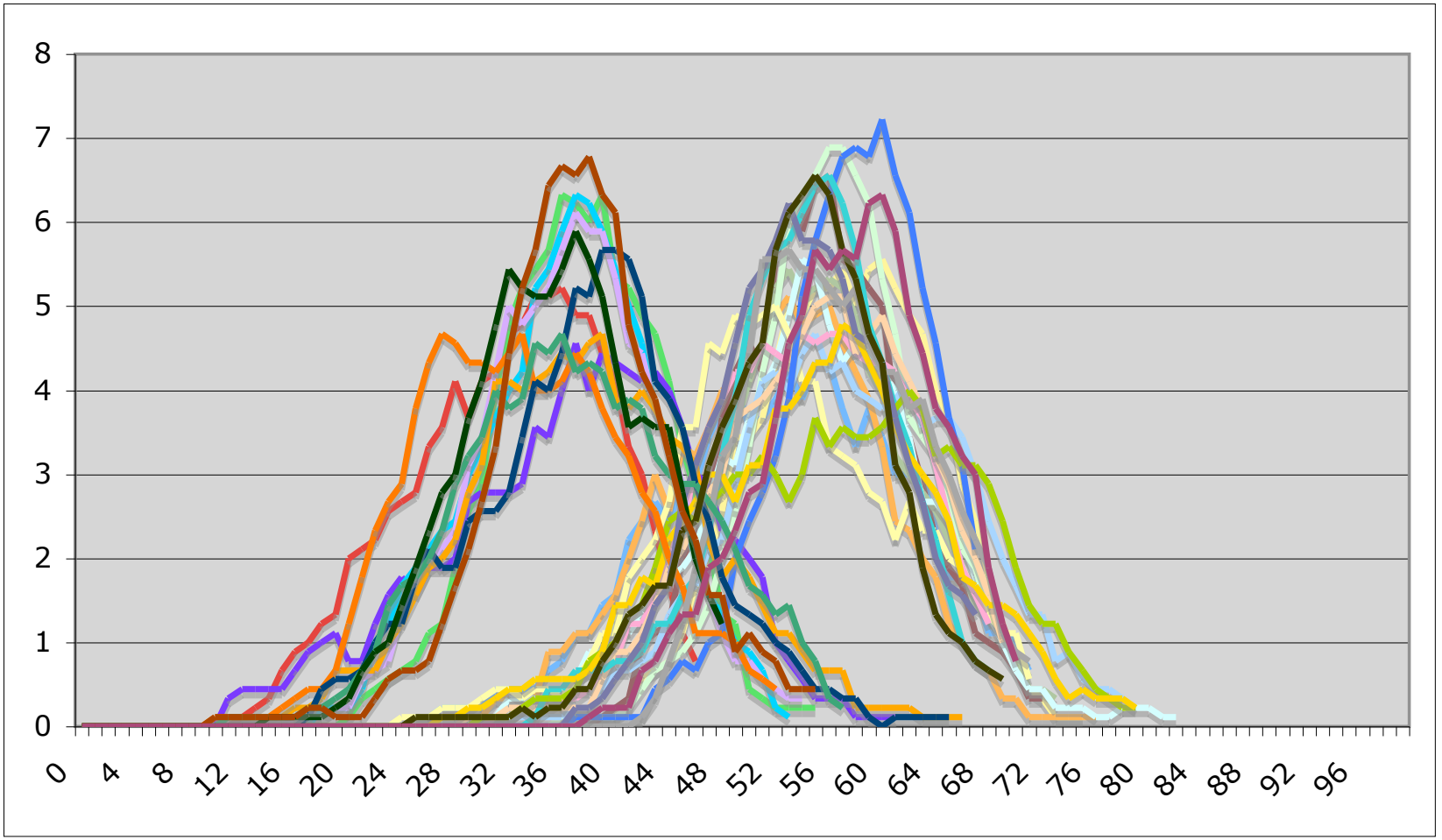
Recall that lenition biases sound outputs toward the center

- If a sound contrast is redundant in a given lexical item, it might be expected to decay toward the center.
  - Might expect just the minimum contrastive sounds per word, with the rest of sounds decaying to neutral.

All C distributions



All V distributions



Summary: language

Given:

1. cross-referenced variation at lexical and sublexical levels
2. reversion to the mean of categories

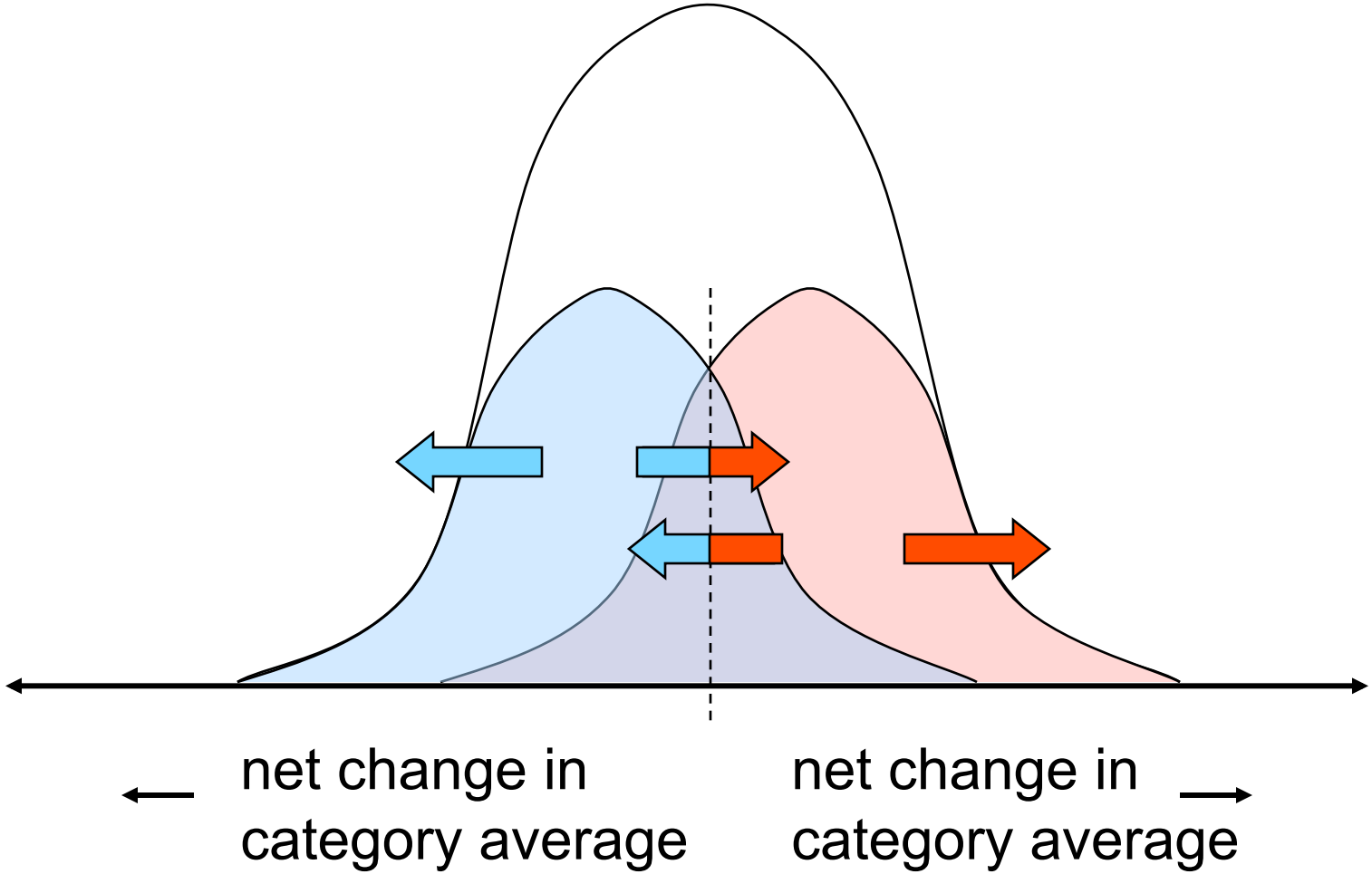
- Any selection for contrast between words promotes the evolution of a coherent, contrastive set of sublexical categories.

Summary: evolution

- Model has no discrete replicators
  - Every exemplar in the model population is a parent to every output.
- Variation is not random: selection acts through biases in which variants arise.
  - Every output is stored in a listener lexical category.

Thank you!

Variation plus competition pulls category means apart



Linguistic memory contains populations of variants.

- Sensitivity to fine within-category variation
  - Exemplar literature (e.g., Johnson 1997)
- Sensitivity to multiple potentially overlapping generalizations
  - Analogical modeling literature (e.g., Skousen 1989, Krott et al. 2001, Ernestus and Baayen 2003, etc.)
  - Studies showing sensitivity to both broad patterns and specific details (e.g., Long and Almor 2000, Kuehne et al. 2000, Albright and Hayes 2002, reviewed in Bybee and McClelland 2005, Pierrehumbert 2007).
- Evidence for gradient change both at sound and word level (reviewed in Bybee 2002).

Fine variant properties are transmitted and reproduced in use

- During acquisition (e.g., Pierrehumbert 2002),
- But also in adulthood (e.g., Goldinger 2000, Harrington et al. 2000)
  - Transmission loops operate at multiple time scales.