

Sublexical contrast as an emergent effect of lexical category competition

Andy Wedel

Dept. of Linguistics, University of Arizona

Pattern formation in complex systems

- Long-range structure often develops in systems that have many repeatedly interacting elements:
 - Interacting positive and negative feedback loops
 - Interactions across levels of organization

Roadmap

1. Category interaction can promote category contrast.
 - Illustration of the interaction of feedback loops to promote structure
2. Given rich categories, promotion of contrast at the lexical level can indirectly promote contrast at the sublexical level.
 - Illustration of cross-level interaction

Two functionally distinct kinds of sound categories

1. Saussurean signs: sound maps to meaning
 - morphemes, words, collocations
2. Sublexical categories: sound is not directly linked to meaning
 - features, allophones, phonemes, syllables

Contrast in Sound Categories

- For lexical categories to be contrastive, sublexical categories must be contrastive.
- Standard models create an information bottleneck between levels:
 - Lexical categories are linked to detailed representations of sound only through the lowest sublexical level.
 - ‘Linkage problem’: how does function of contrast at the lexical level influence trajectories of sub-categorical change at the sound level?

Questions

1. How do these contrastive sound categories come about?
2. 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.
 - Vowel inventory structure
 - Chain shifts
 - Contrast maintenance (homophony avoidance) in paradigms

Language as a system supports evolutionary pathways of change

- Populations of variants at multiple levels
- Transmission/propagation through use
 - Both at the population/generational levels, but also within individuals and generations.

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).
- Variation recorded at multiple levels of organization (reviewed in Bybee 2002, Pierrehumbert 2007)

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.

Elements I'll use here:

1. Categories at all levels retain some record of variation.
2. Both perception and production influenced by category contents:
 - perception/production feedback loop

Arguments

- Contrast maintenance through the course of sound-change *does not* require a teleological mechanism.
 1. The statistics of mapping material into and out of competing categories promotes contrast between those categories.
 2. Cross-referenced variation at multiple levels of organization opens up the possibility that contrast maintenance arises *interior* to the chain of nested categories, e.g., at the lexical category level rather than at a sublexical level.

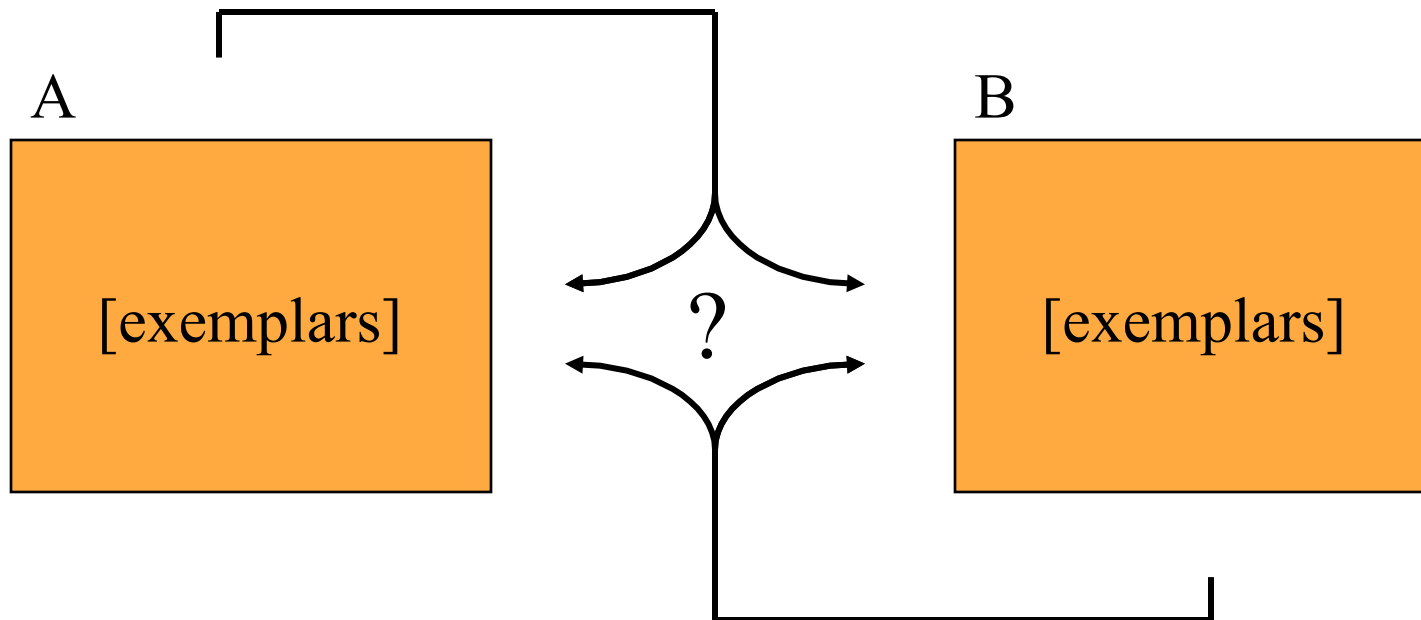
Part 1: Category contrast through competition

- Mapping of percepts to categories proceeds through a 'competition' between categories for assignment to the percept (e.g, NAM, Luce and Pisoni 1998).
- If categories compete on the basis of their contents, this results in segregation of distinct percept types to distinct categories.

Exemplar-based simulation

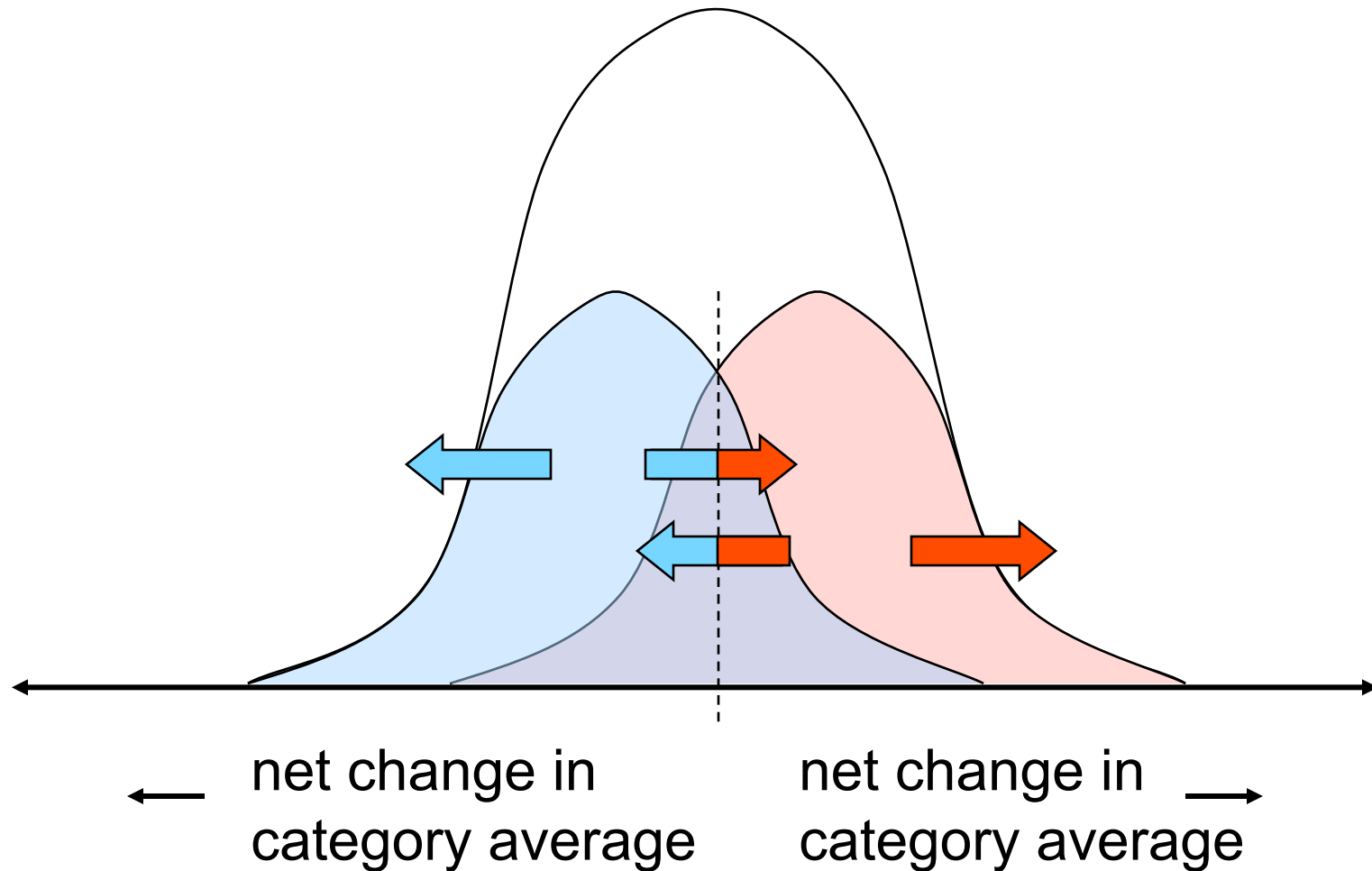
- Two categories, each containing numerical exemplars.
- Produce outputs from each category
 - Production with noisy reversion to the mean of the category

Experimenters test as a new exemplar in the same category with the closest mean.



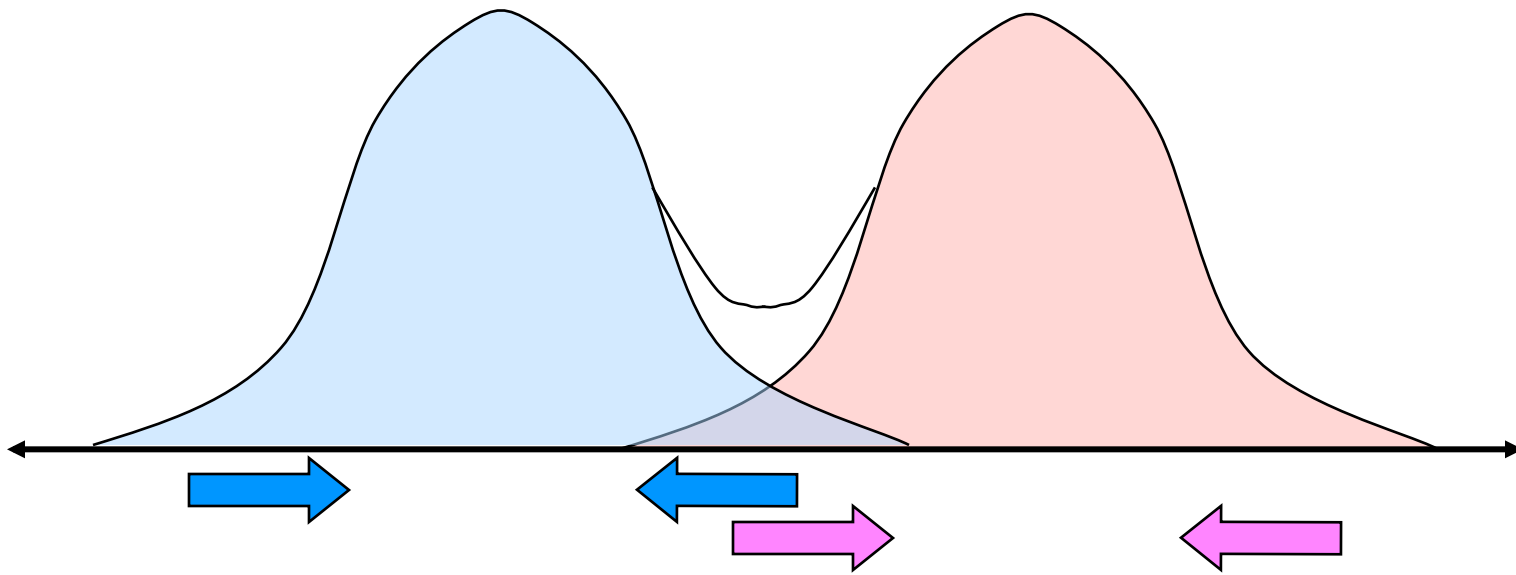
Category Competition Simulation

Variation plus competition pulls category means apart



Non-discreteness of representations creates reversion to the category mean

- Blending in production or perception pulls category boundaries back toward the mean.
- The interplay between variation and bias to the mean creates distinct distributions.

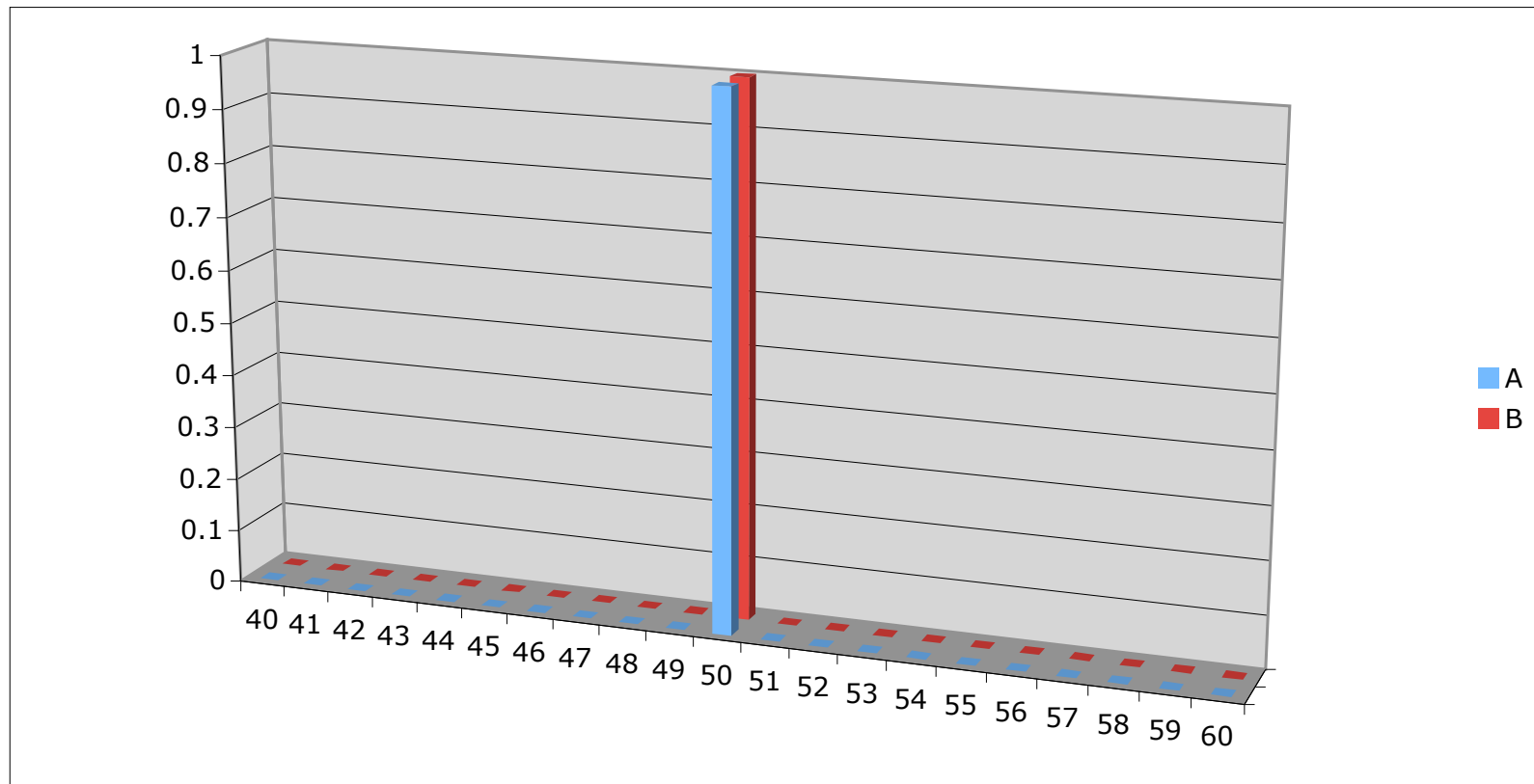


Example with an alternative architecture: connection weights

- Define a perceptual dimension from 1-100.
- Each category has connections to each point with weights ranging from 0-1.
- Weights are incremented with use, decremented with non-use.
- Activity is relative to weight: 'priming' creates a positive feedback between past and future use.

Initial state

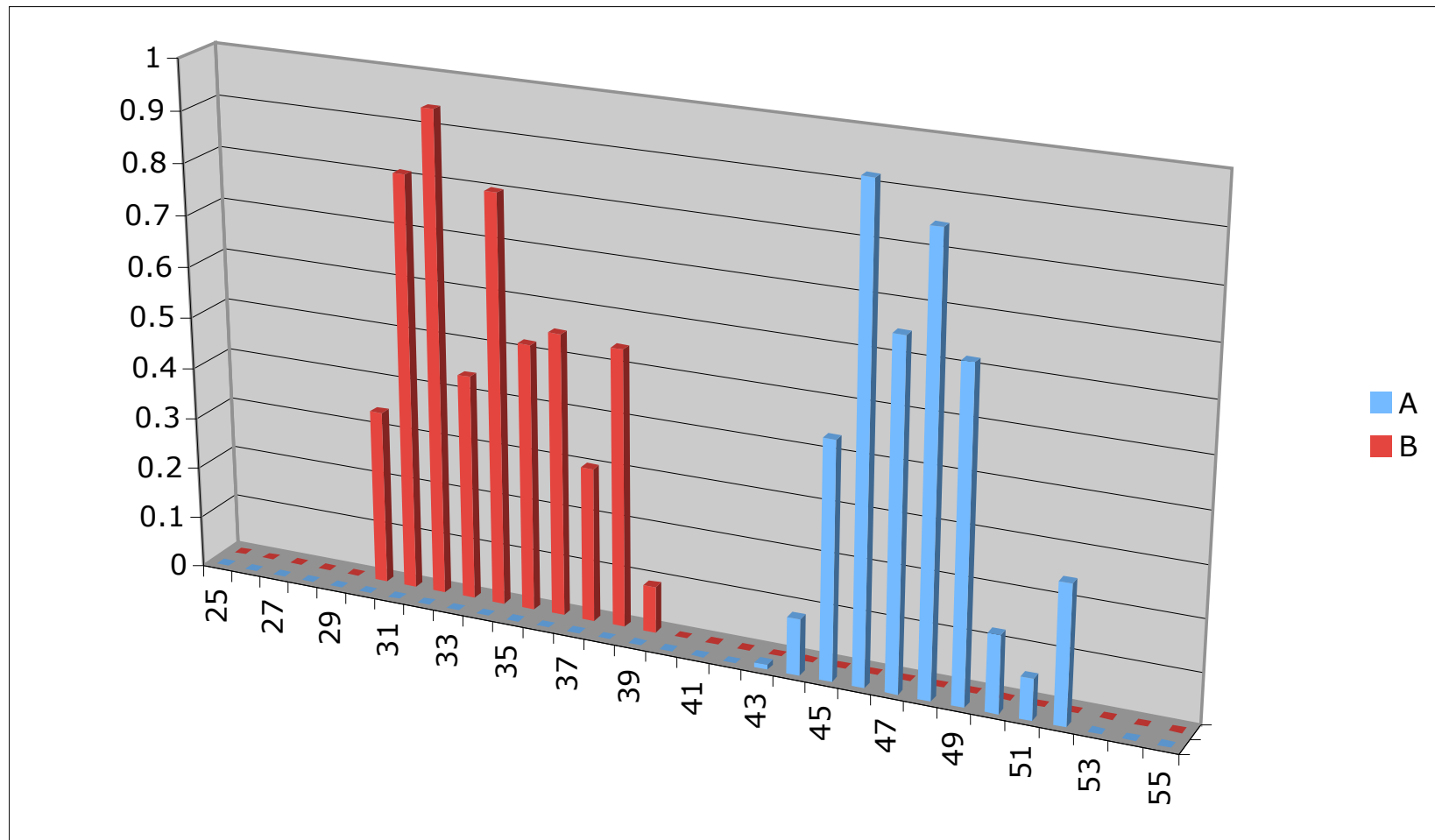
- Start with single fully active weights at 50 in both categories



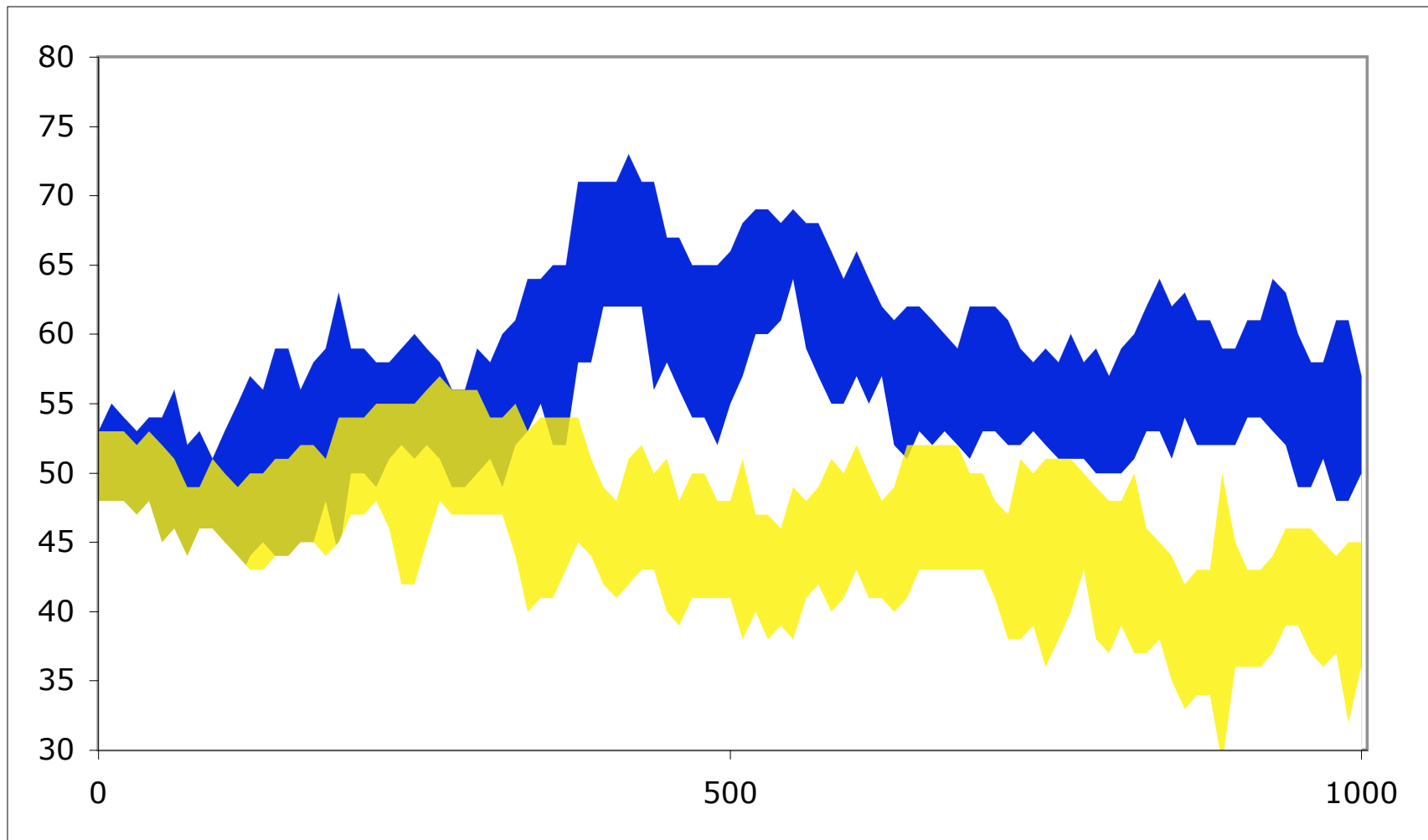
Production/categorization loop

- In each round:
 - 1) Produce an averaged output from each category.
 - Output biased toward the center of the distribution
 - 2) Compare the output to the contents of the two categories.
 - 3) Probabilistically store the output in the closer category.
 - Increase the weight of the link to the jnd unit as the inverse of the exponential decay rate.
 - 4) Memory decay
 - Exponentially decrement all activations.

Category contents at 2000 rounds

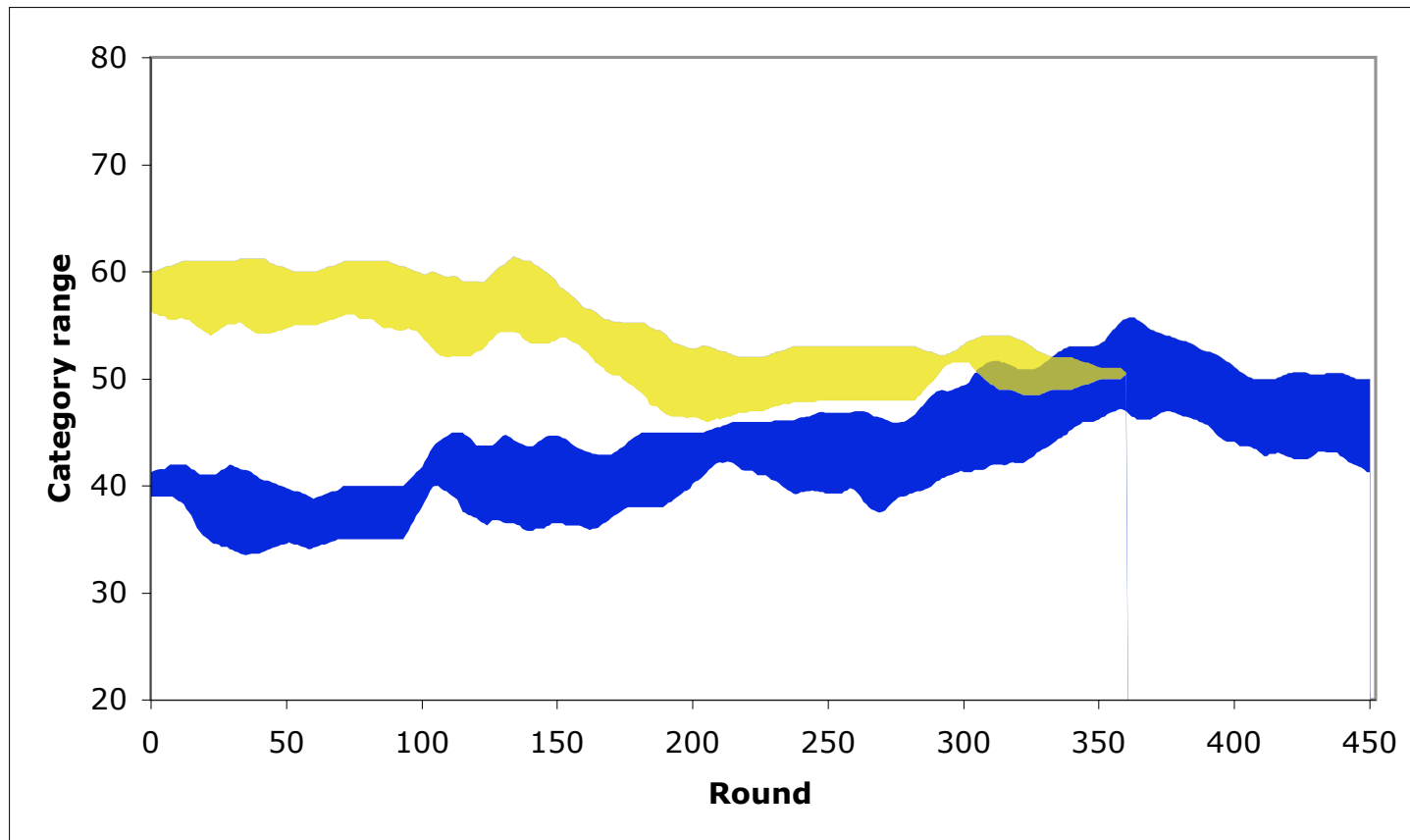


Change in distribution over time



Control 2: no independent anchoring of categories

- Probability of production from a category is in relation to the total category activation (Pierrehumbert 2002).



Important conclusion

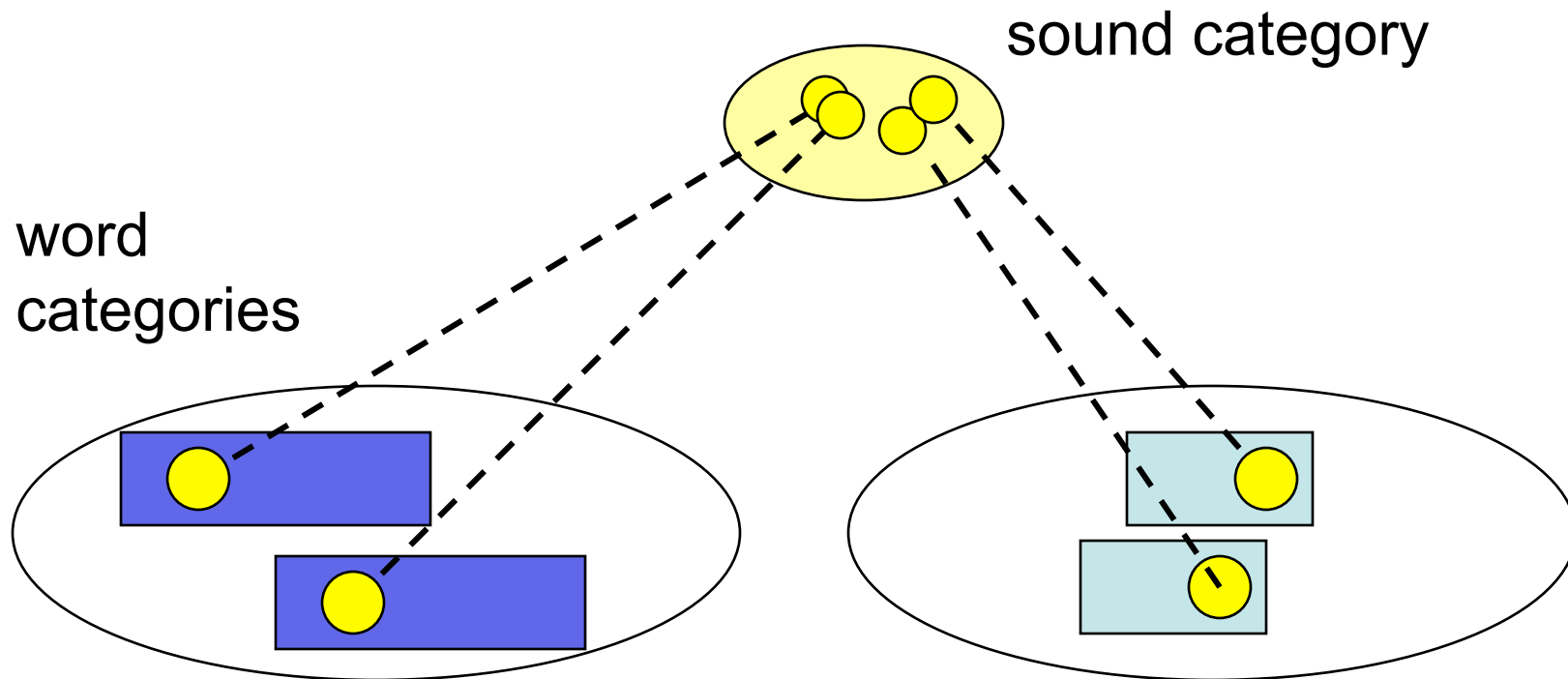
- To be stable in the face of competition, categories must be anchored independently of their activations, that is, their history of use.
- Sound/meaning categories fit this bill.

Part II: Cross-level interaction

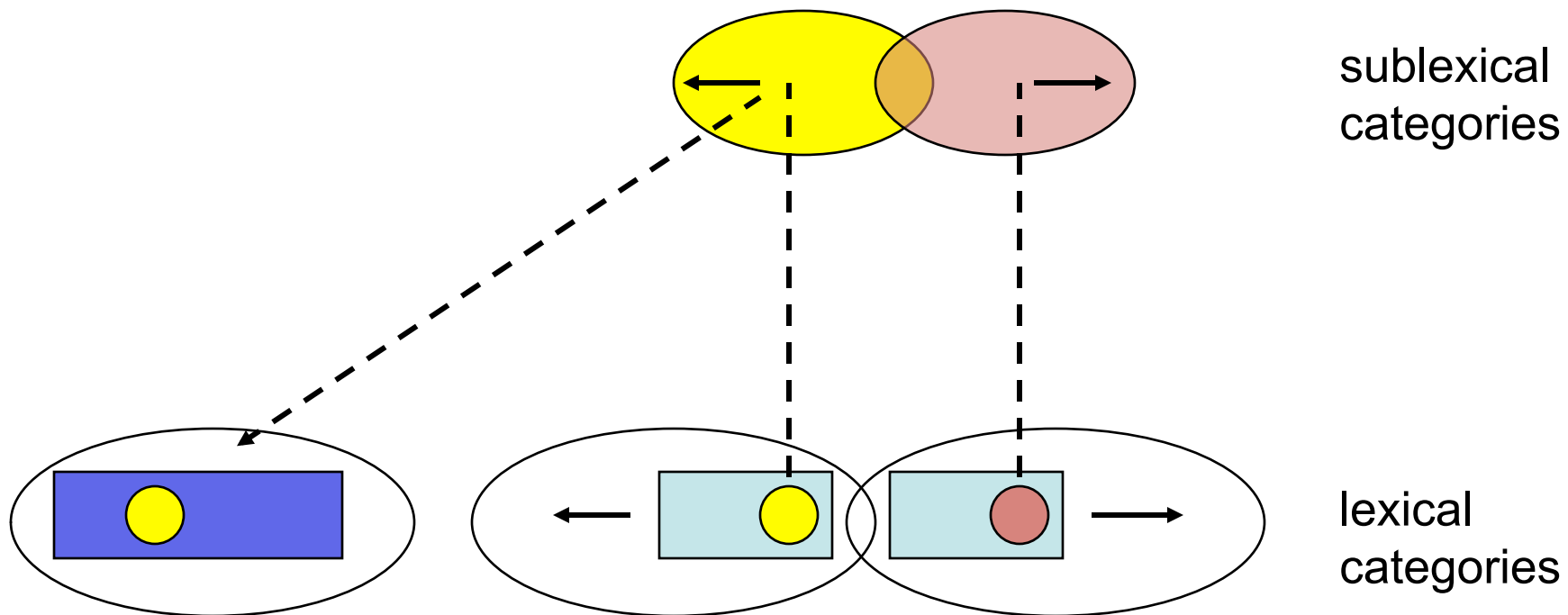
- Most important conclusion I want to make:
 - Contrast maintenance at the word level can create contrast at the sound level.
 - Why is this relevant?
 - => Sound-meaning category production is independently anchored in use, sound category production is not.

Model set-up

- Words are composed of an ordered set of sounds.
- Sound and word categories consist of cross-referenced exemplars (e.g., Bybee 2002).
- 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.
 - Selection is context-dependent
- Over time, contextually adaptive gene variants spread through the population.

Model set-up

- Sound memories are represented as activated positions in a 20x20 grid.
- Two kinds of categories:
 - Lexical
 - The category label is a meaning and an ordered pair of sound category labels.
 - Category contents consist of a set of word exemplars with associated activations
 - Sound
 - The set of category labels is defined in relation to the distributions of activated positions in the grid.
 - Category contents consist of a set of corresponding sounds taken from words containing that label, with the associated activation.

Cycle

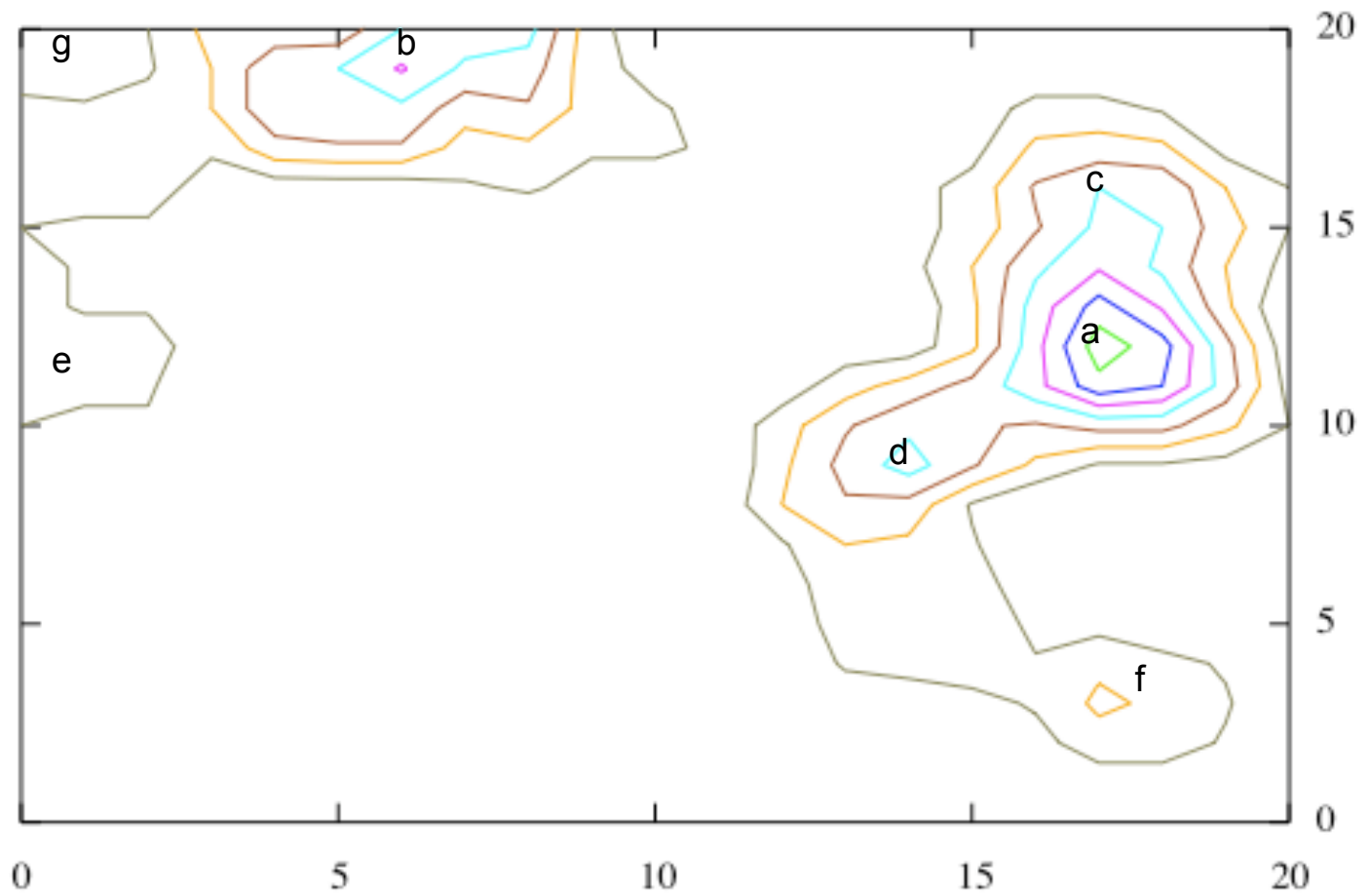
- Production:
 - randomly choose three word exemplars from the word category
 - Output is an average of the word exemplars, with some bias from the corresponding sound categories.
 - Slight leniting bias at the sound level toward the center of the grid.
- Perception:
 - listener compares output to all word exemplars in its lexicon
 - stores the output as a new exemplar in the best matching lexical category.
- Category competition is solely at the word-level
 - Purposeful heuristic feature: If sound category contrast develops, it can only do so indirectly via maintenance of word category contrast.

Generational change

- *Within* a generation, multiple agents talk back and forth, producing outputs, categorizing percepts and storing new exemplars.
- Competition is between word-categories only
 - Heuristic restriction!
- *Across* generations, word category labels and contents are transmitted to new agents, but not sound categories.
- Each new generation deduces sound categories anew from the acquired distribution of sounds in word exemplars.
 - Use a QT clustering algorithm (Heyer et al, 1999)

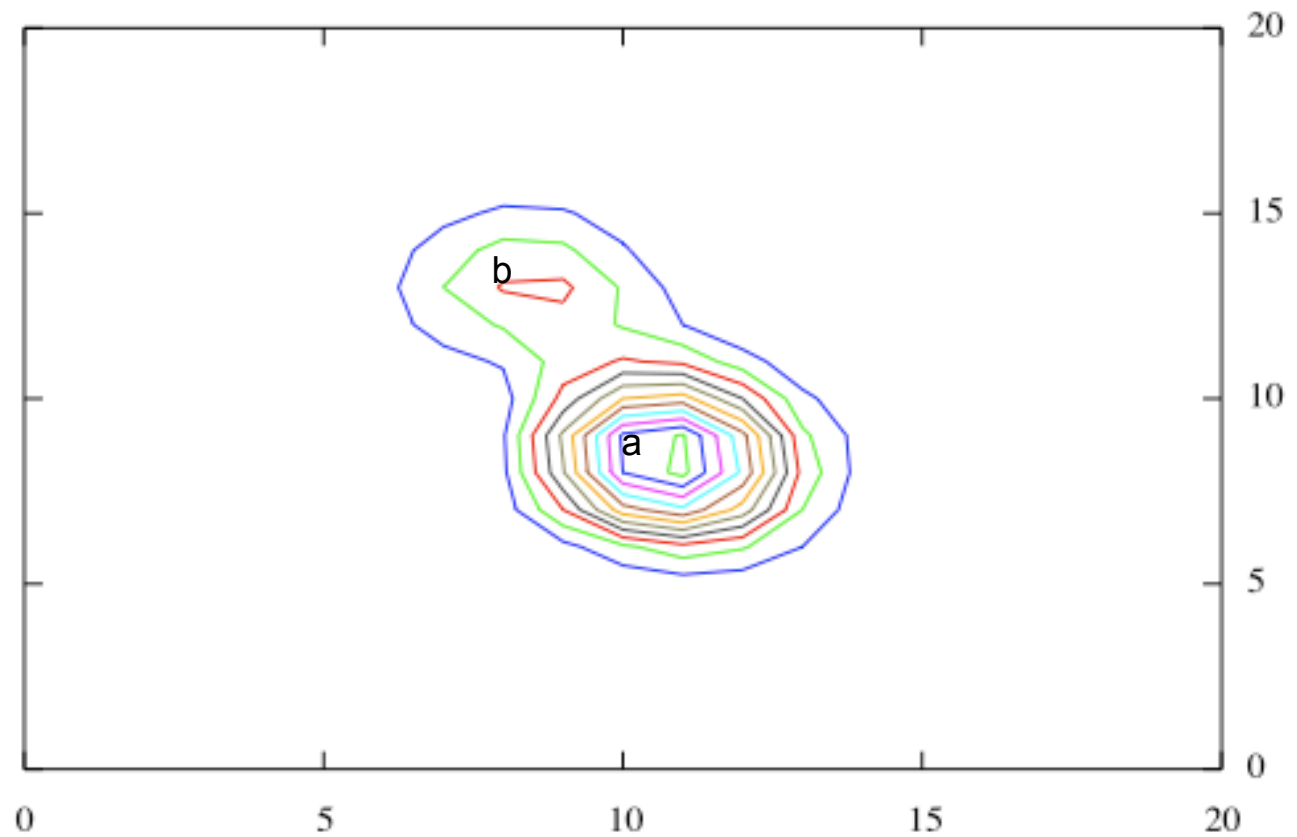
Control: No category competition, round 50

<u>Word</u>	<u>Sounds</u>
1	c b
2	b a
3	c e
4	a c
5	e a
6	b f
7	d b
8	a a
9	a g
10	c d



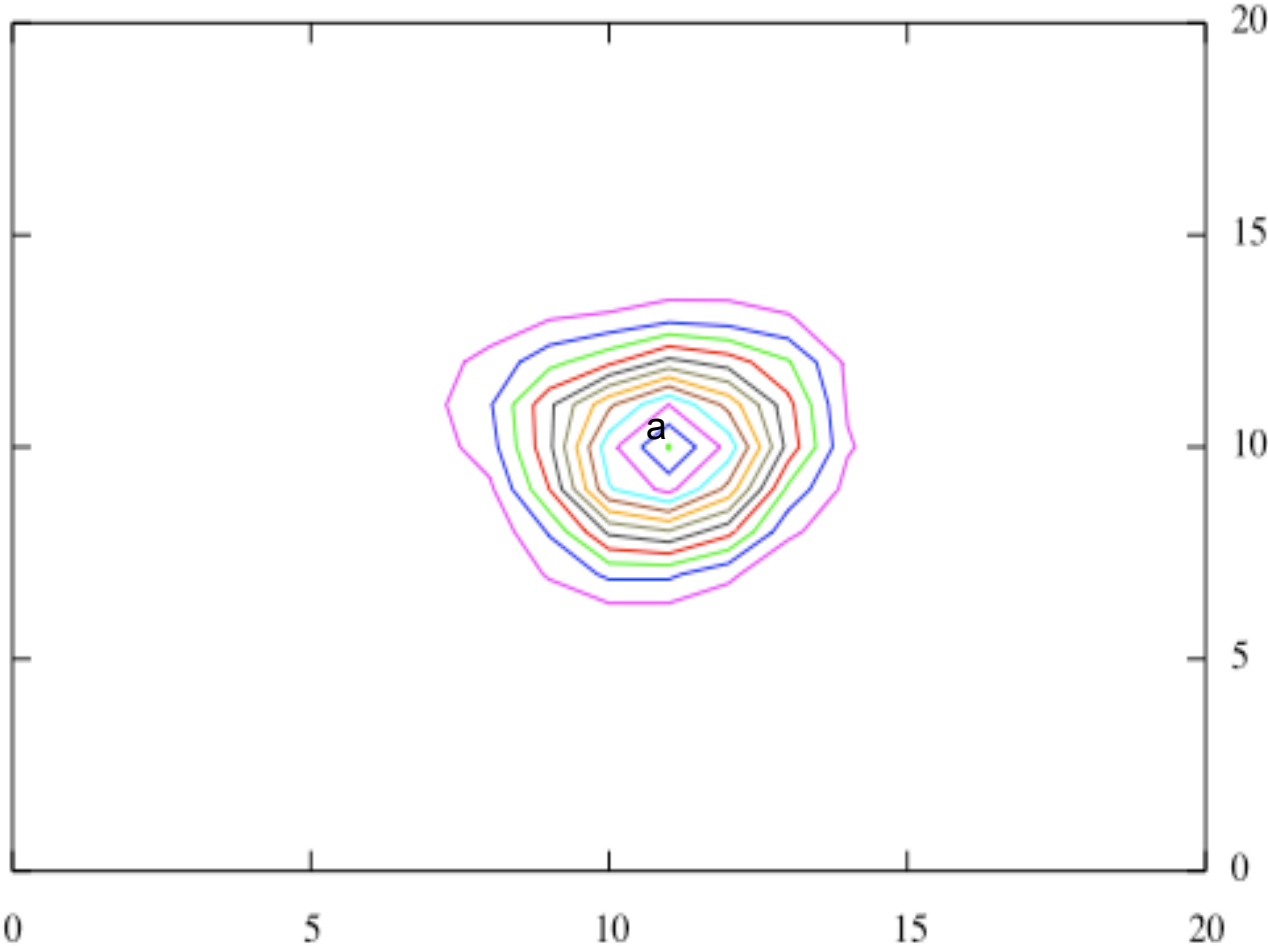
Control: round 1000

<u>Word</u>	<u>Sound</u>
1	a a
2	a a
3	a a
4	a b
5	a a
6	a b
7	a a
8	a a
9	a a
10	b a



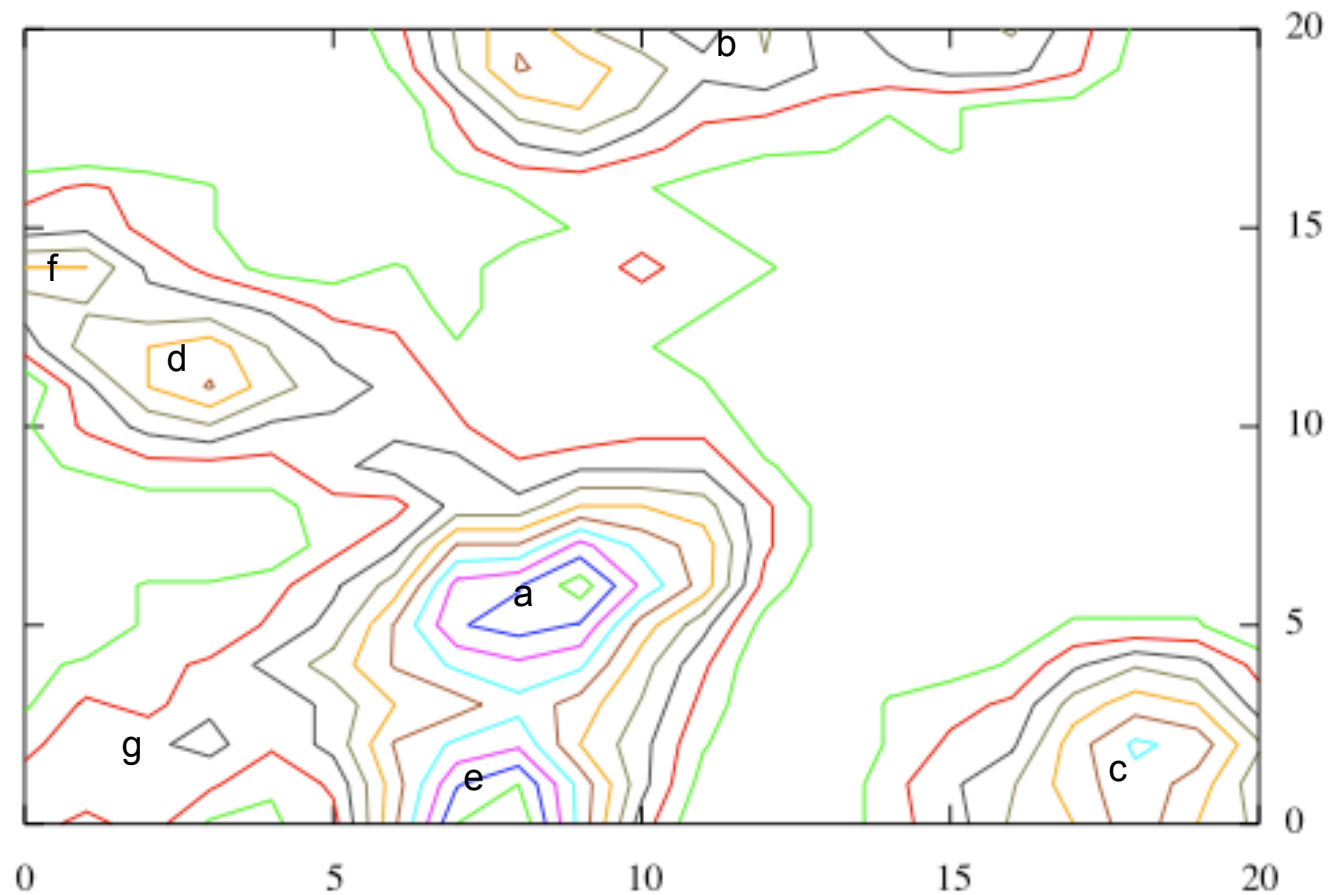
Control: round 2000

<u>Word</u>	<u>Sound</u>
1	a a
2	a a
3	a a
4	a a
5	a a
6	a a
7	a a
8	a a
9	a a
10	a a



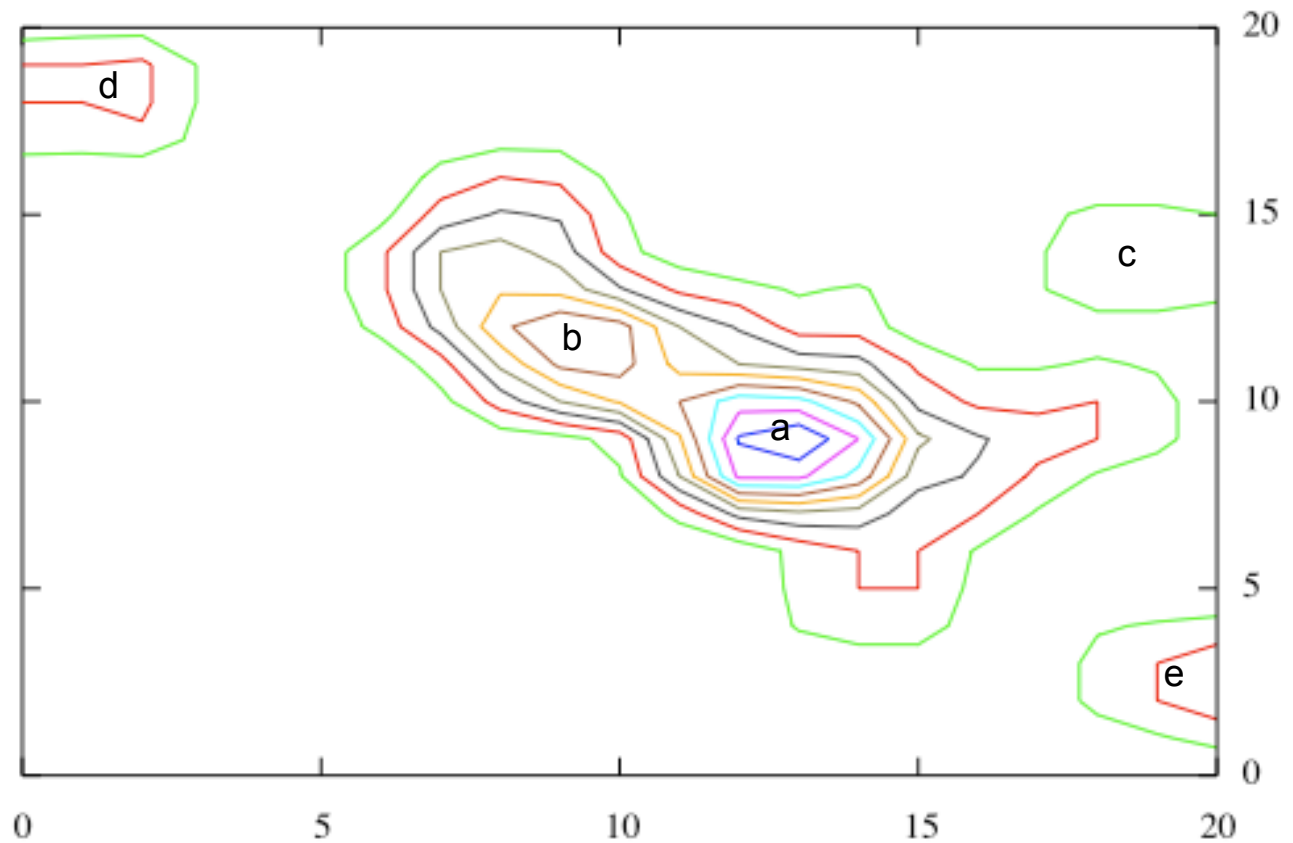
Category Competition: 50

<u>Word</u>	<u>Sounds</u>
1	a f
2	g d
3	a e
4	b c
5	a e
6	f b
7	a a
8	c d
9	a b
10	b c



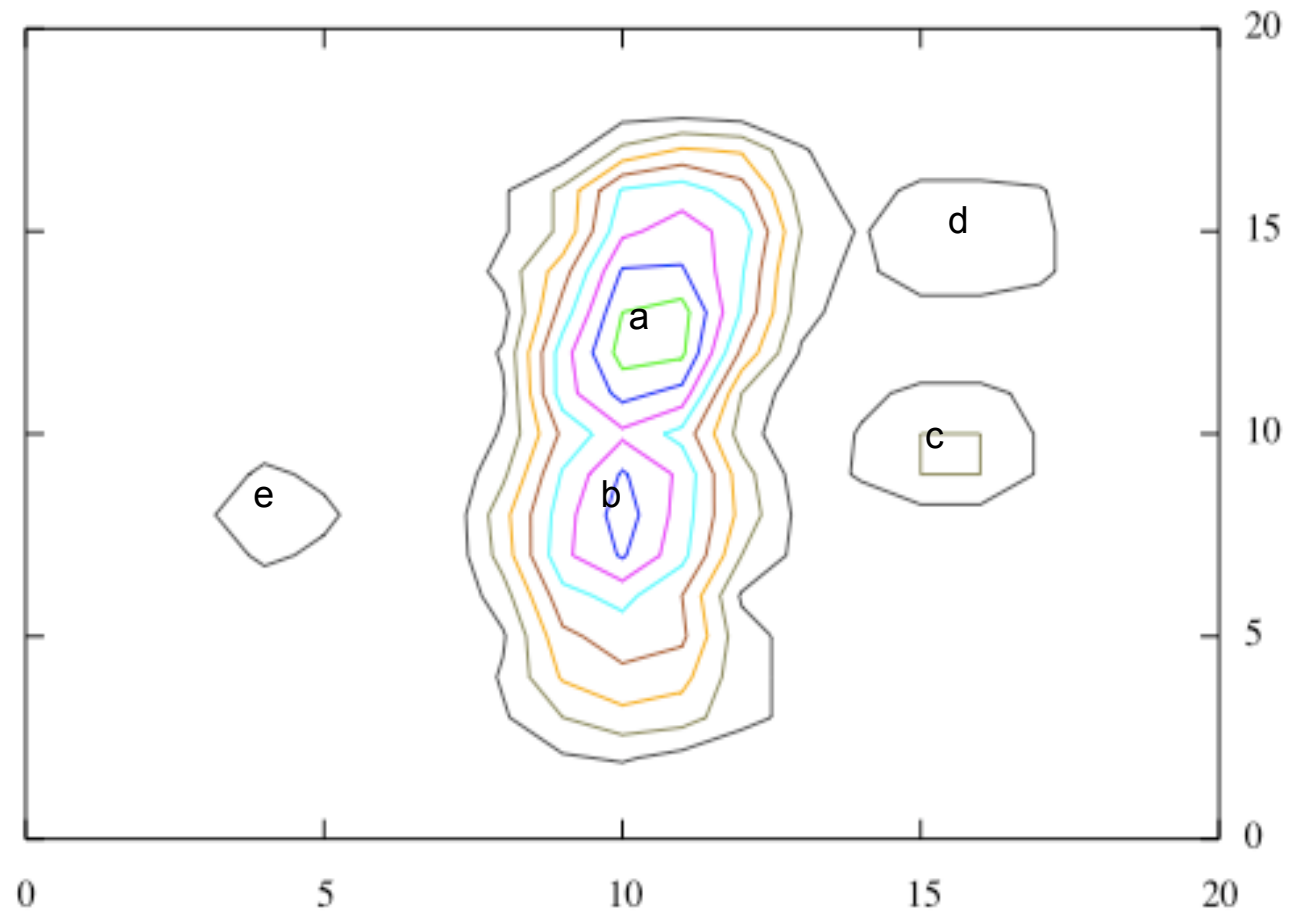
Category competition: 2000

<u>Word</u>	<u>Sounds</u>
1	a b
2	b d
3	a a
4	e a
5	a a
6	b c
7	a b
8	b a
9	b b
10	d b



Category competition: 4000

<u>Word</u>	<u>Sound</u>
1	a a
2	b b
3	a b
4	a a
5	b a
6	d a
7	a b
8	c a
9	e a
10	a b



Conclusions

1. Given cross-categorization of sound variants at the sound and word level, any processes that keep words contrastive indirectly keep sounds contrastive.
2. Contrast in phoneme inventories may be an epiphenomenon of contrast-promoting processes at higher levels.
 - An example of mutual influence between two levels of organization.

Part I and II conclusions are independent

- Other, non-exclusive possibilities beyond category competition exist for the support of contrast.
 1. Selection against ambiguous forms in categorization
 2. Hyperarticulation of potentially ambiguous forms.
- Any and all of these may operate at the word level and feed back to the sound level.

Further work

- Simulations allow one to investigate models with precision, but they cannot constitute proof that a given system operates in a particular way.
- Need back and forth bootstrapping between modeling and experiment.
 - Evidence from anti-homophony effects in paradigms (Blevins).
 - Are existing homophones semantically more distant than average word pairs?

Thank you!