Bilinguals' categorization behavior: Internally or Externally Emergent?

18 June 2011

Benjamin Zinszer & Ping Li
Pennsylvania State University

Brain, Language, and Computation Lab
Center for Language Science

Lexical Categorization (English Monolingual)

Emergent Bilingual Categorization

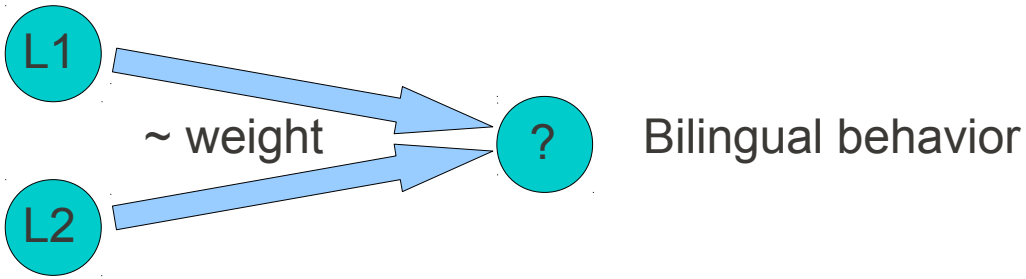
- “Similar” objects are frequently grouped together under a common label (lexical categorization)
- Lexical categorization presents a set of associative conflicts to bilinguals as languages do not categorize universally.

How do bilinguals resolve conflict?

- A few possibilities:
  - Bilinguals maintain exactly monolingual-like representations in each language
  - Bilinguals choose a single language to set category boundaries
  - Bilinguals compromise between language systems where conflicts arise

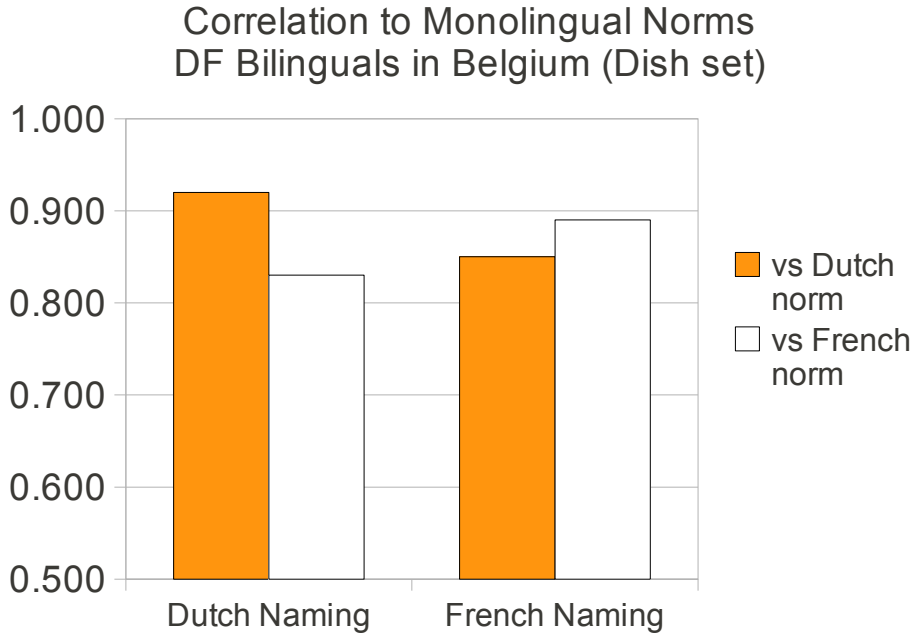
What about associative learning?

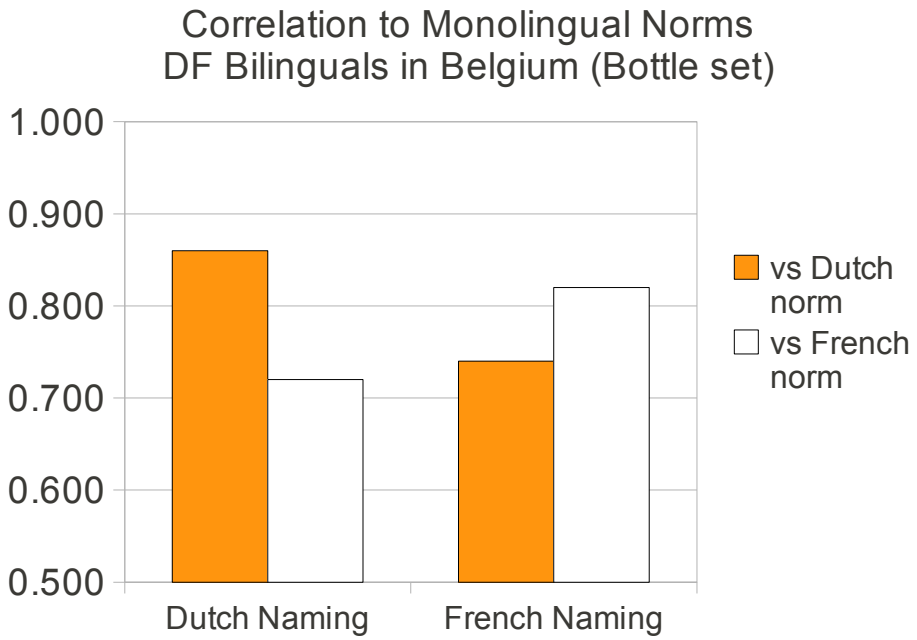
- Tendencies of associative systems:
  - Two conflicting naming systems would be in constant competition for activation, leading to unstable production pathways
  - A single monolingual-like representation would be extremely difficult to maintain in the face of constant conflicting input
  - Preference for language-specific patterns modulated longitudinally by language input



Empirical Evidence from Simultaneous Bilinguals

- Dutch-French simultaneous bilinguals show convergence between languages in categorization patterns. (Ameel et al, 2005)





Cross-Language Correlations:

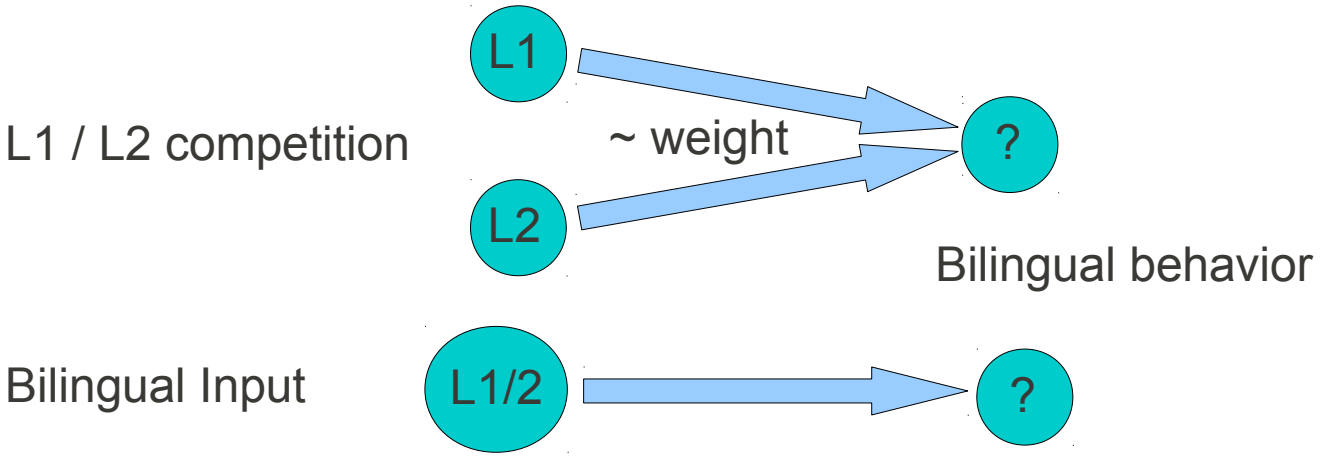
Mono-mono: 0.80 < Biling-biling: 0.91

Cross-Language Correlations:

Mono-mono: 0.63 < Biling-biling: 0.88

Question of Emergence...

- Does convergent naming system arise socio-linguistically in the bilingual community or as a competitive learning process?
- It is difficult to tease apart the effects of these two forces in a bilingual community.



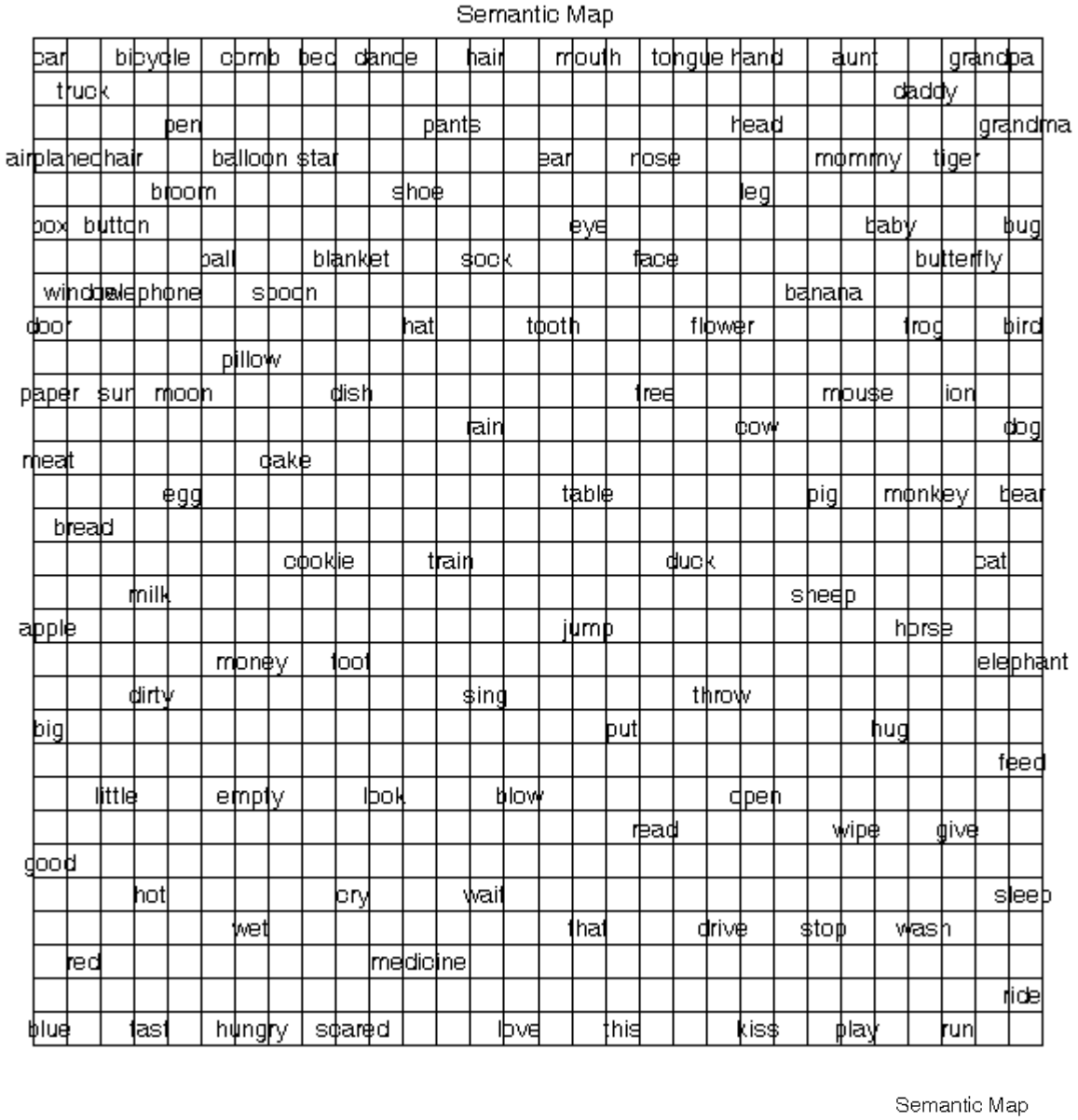
Modeling monolingual inputs

- Dominant monolingual names in Chinese and English for 73 containers → PatPho phonological coding system
- Semantic ratings of each container, e.g.:
  - “It has a straw”
  - “It contains a liquid”
  - “You can cook in it”

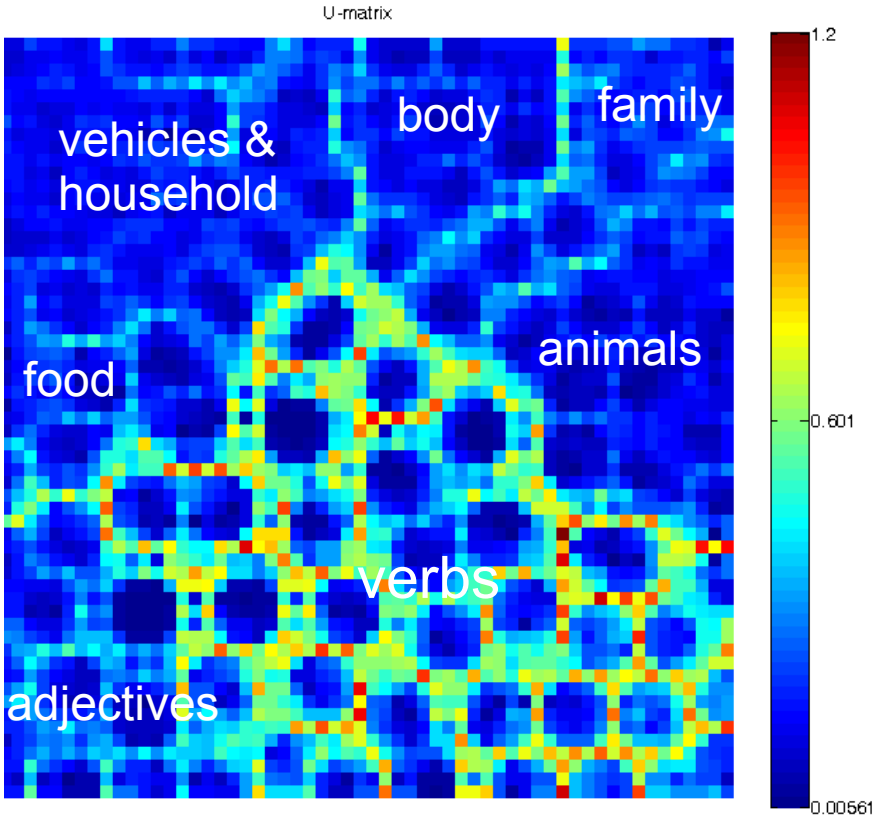
… vectorized by rater agreement: [1.0 0.8 0.0]

Monolingual naming norms and semantic ratings provided by Malt

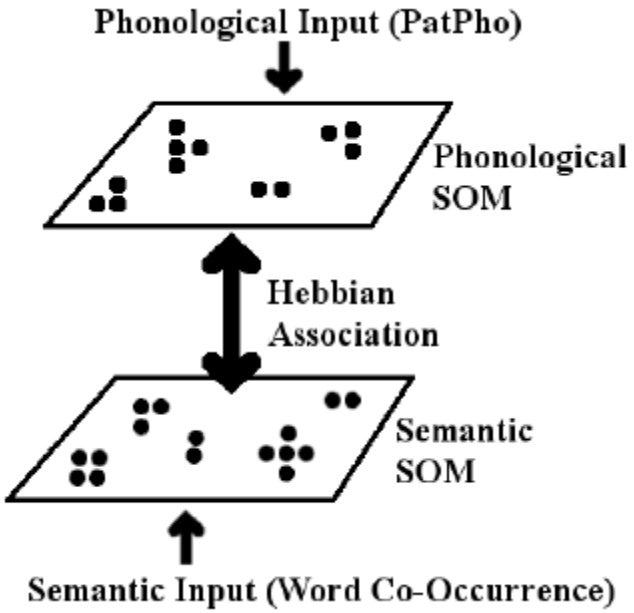
Self Organizing Feature Maps (SOM)



Semantic Map



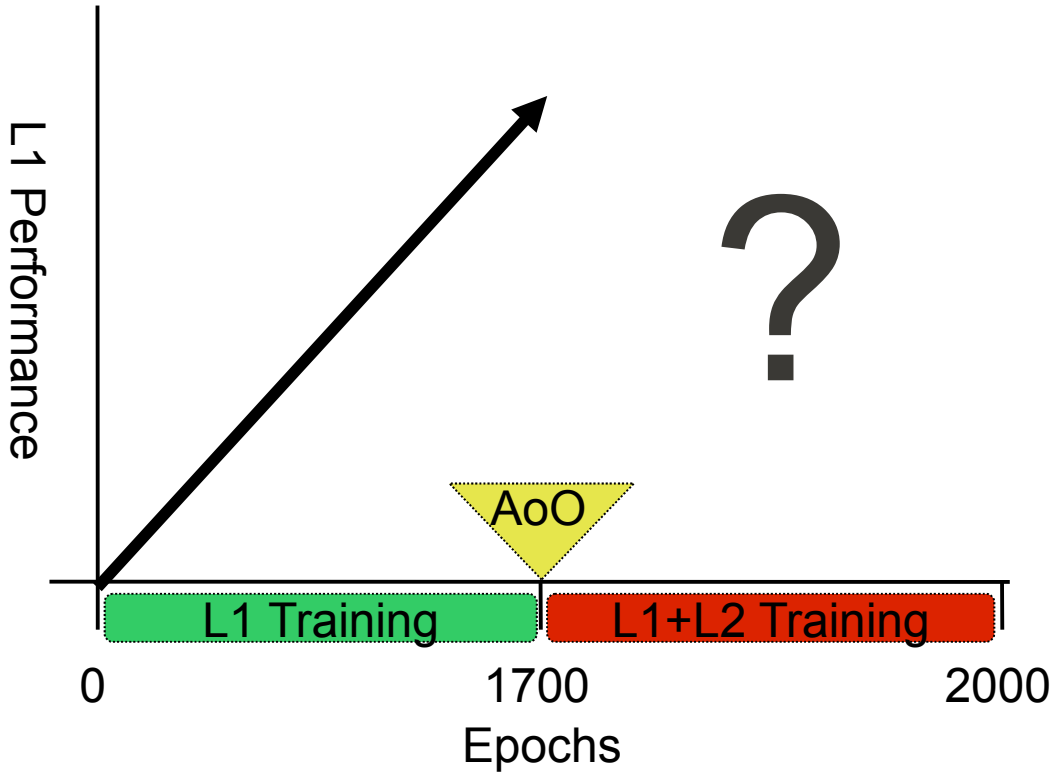
The Model – Architecture & Parameters



- Phonological [15x15] & Semantic [25x25] SOMs
  - Gaussian neighborhoods
  - Dynamically decreasing radius (13 to 1)
  - Alpha 0.2 (constant)
- Two sets of saturating Hebbian connections
  - Every S-node to every P-node
  - Language selection nodes to every P-nodes
  - Beta 0.2 (constant)

The Model – Training

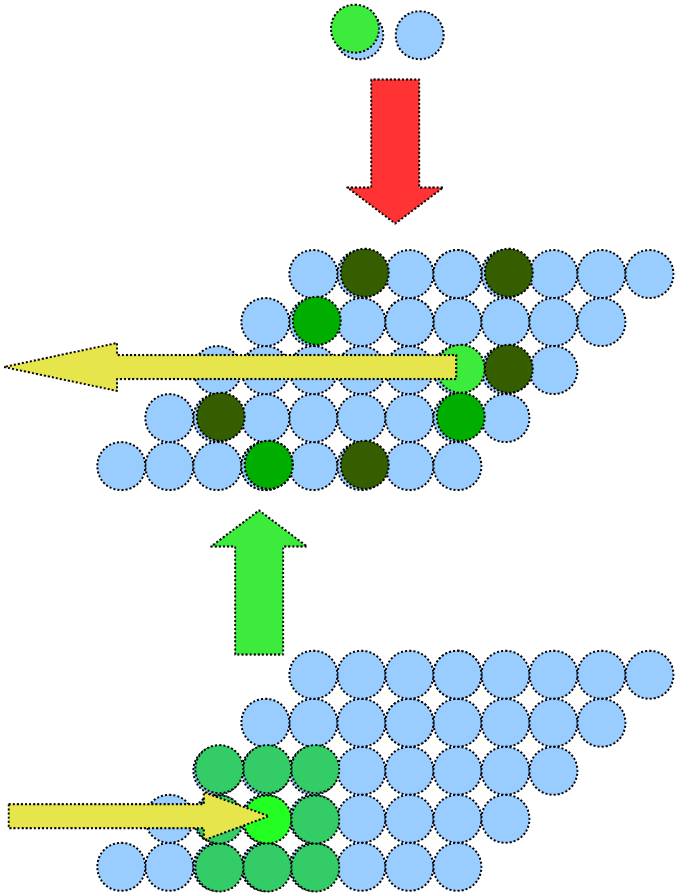
- Stimuli – 73 containers, named by monolinguals in E and C
  - 42-dimensional phonological vector (PatPho)
  - 68-dimensional semantic vector (feature ratings)



- Age of Onset (AoO) 1700 epochs
- 300 epochs of varying L2 input after AoO

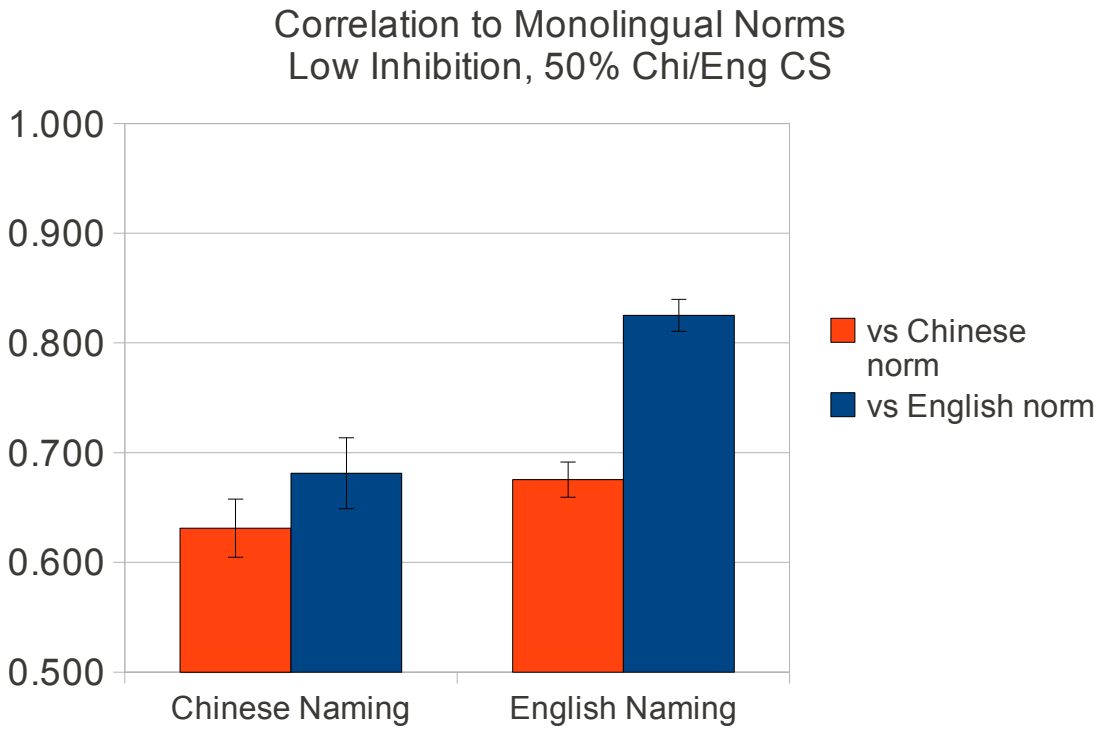
The Model – Testing

- Semantic map and language selection nodes activated with a stimulus
- Activation propagated via Hebbian connections
- Activation measured on phonological map
- Activation patterns compared to BMUs



10 Simultaneous CE Bilingual Models

- English dominance in naming patterns
- Cross-linguistic correlations differ from DF bilinguals
- BUT, within-bilingual correlations higher than between monolingual, suggesting L1-L2 convergence

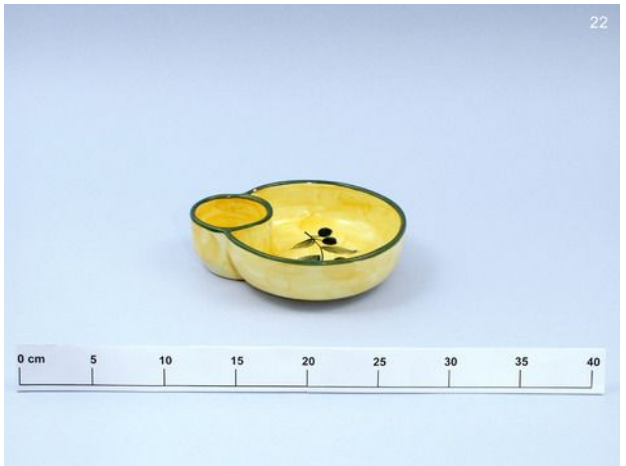


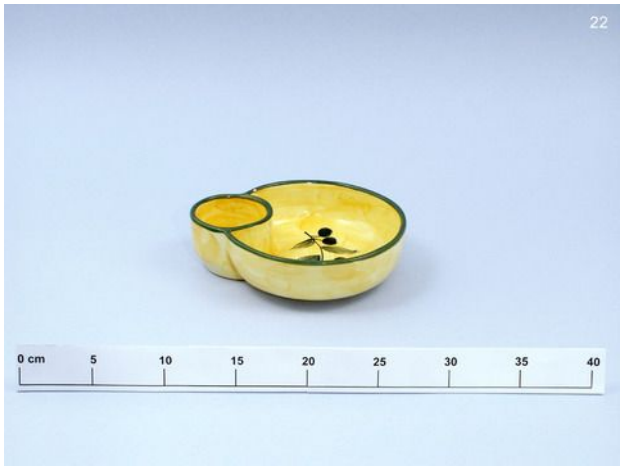
Cross-Language Correlations:

Mono-mono: 0.57 < Biling-biling: 0.69 (+/- 0.07)

C-E Bilingual Humans

	Chinese means & sds		English means & sds	
n	12		12	
Age	19.92	1.31	19.75	1.36
AoA	17.0	3.02	16.5	2.81
LoR	3.17	3.13	3.50	3.03
TOEFL	91.3	8.5	90.8	8.1
CS Freq (max 10)	3.25	2.83	3.58	2.64



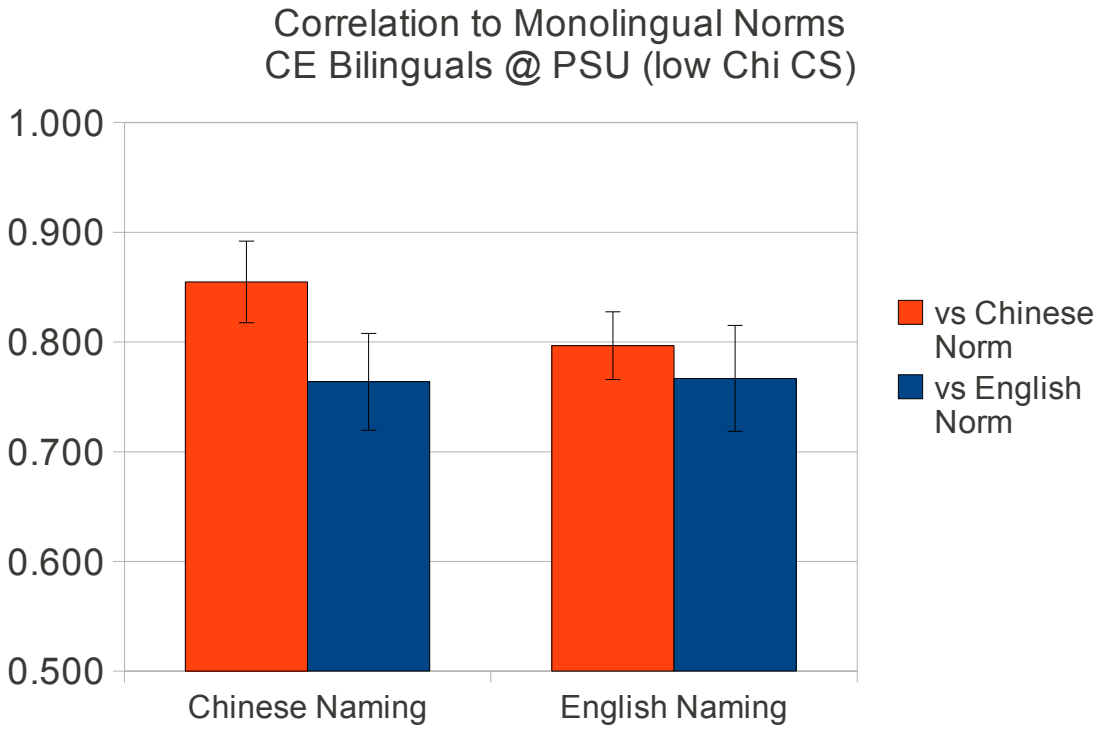


这是什么？

What is this?

Behavioral Results

- Chinese dominance in naming patterns
- Reversed relationship to simulation of simultaneous bilinguals
- AGAIN, within-bilingual correlations higher than between monolingual, probably L1 transfer to L2



Cross-Language Correlations:

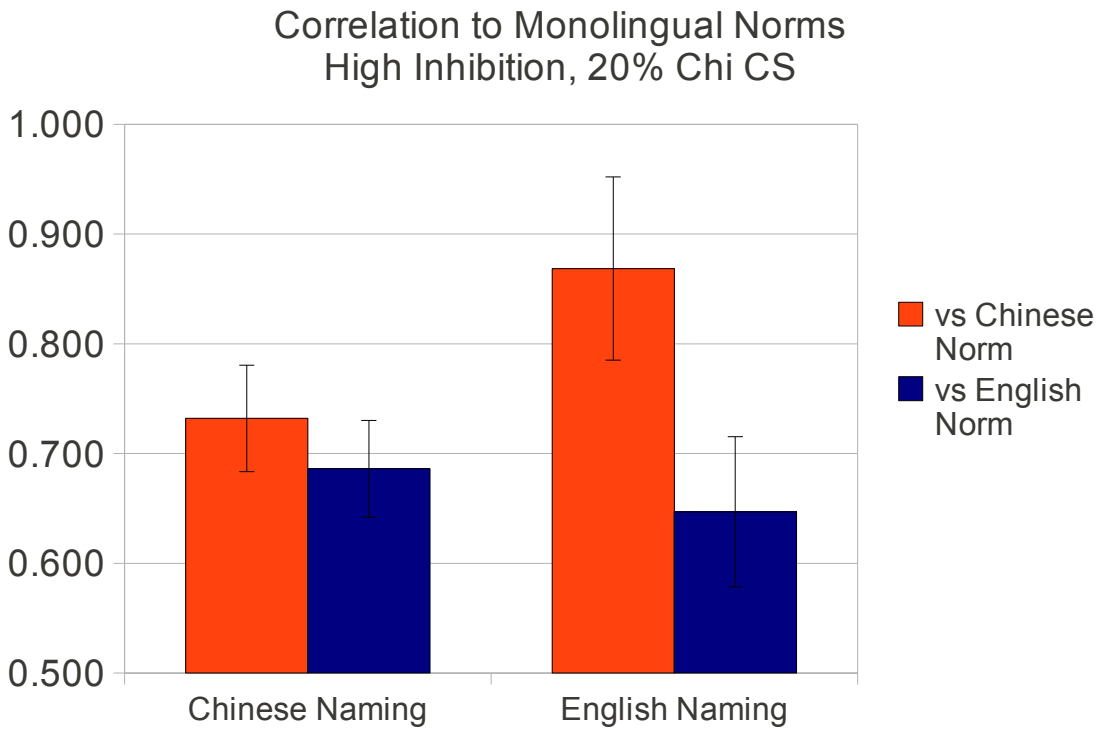
Mono-mono: 0.78 < Biling-biling: 0.81 (+/- 0.02)

Simulating PSU CE Bilinguals

- Set a later AoO in the model (previously 0)
- Low code-switching suggestive of high cross-language inhibition (speakers maintain separate monolingual domains) → increase inhibition
- Set code-switching frequency lower than 0.5 for Chinese
- Using bottle set, instead of dish set

Model Results

- Replicates pattern of Chinese dominance in Chinese and English naming
- Chinese naming also shows effect of L1
- Within-bilingual correlation still higher than between monolingual



Cross-Language Correlations:

Mono-mono: 0.57 < Biling-biling: 0.75 (+/- 0.05)

Summary & Conclusions

- Monolingual input to low-CS bilinguals at early LoR shows L1-to-L2 transfer, increased convergence over monolinguals
- SOM model accurately predicts this behavior in Chinese-English, late AoO bilinguals
- SOM model predicts different simultaneous CE bilingual convergence patterns than observed in human DF bilinguals
- Convergence in bilingual naming patterns from monolingual-like input is computationally plausible, but actual behavior probably also modulated by naming conventions of bilingual community.

Thank you, 谢谢

Andrew Barney
Patrick Clark
Jason Gullifer
Elisabeth Paulson
Sebastian Rotolli
Jing Yang
Dan Zhong

Center for Language Science
National Science Foundation