# A Possible Division of Labor between Arbitrary and Systematic Sound-Meaning Mappings in Language

Morten H. Christiansen
Cornell University
Santa Fe Institute

# The Arbitrariness of the Sign

*"Onomatopoeia and sound symbolism certainly exist, but they are asterisks to the far more important principle of the arbitrary sign—or else we would understand the words in every foreign language instinctively, and never need a dictionary for our own!"*

Pinker (1999: 2)

*"... the particular phonological forms that a language chooses to convey particular concepts [...] generally are truly arbitrary, except in relative rare cases of phonaesthemes."*

Goldberg (2006: 217)

# Universal Language

"And out of the ground the LORD God formed every beast of the field, and every fowl of the air; and brought them unto Adam to see what he would call them: and whatsoever Adam called every living creature, that was the name thereof."

King James Bible, Genesis 2:19

"The Building of the Tower of Babel" by Hendrik Van Cleve
(1525-1589) Kröller-Müller Museum

# Uncovering the Perfect Language

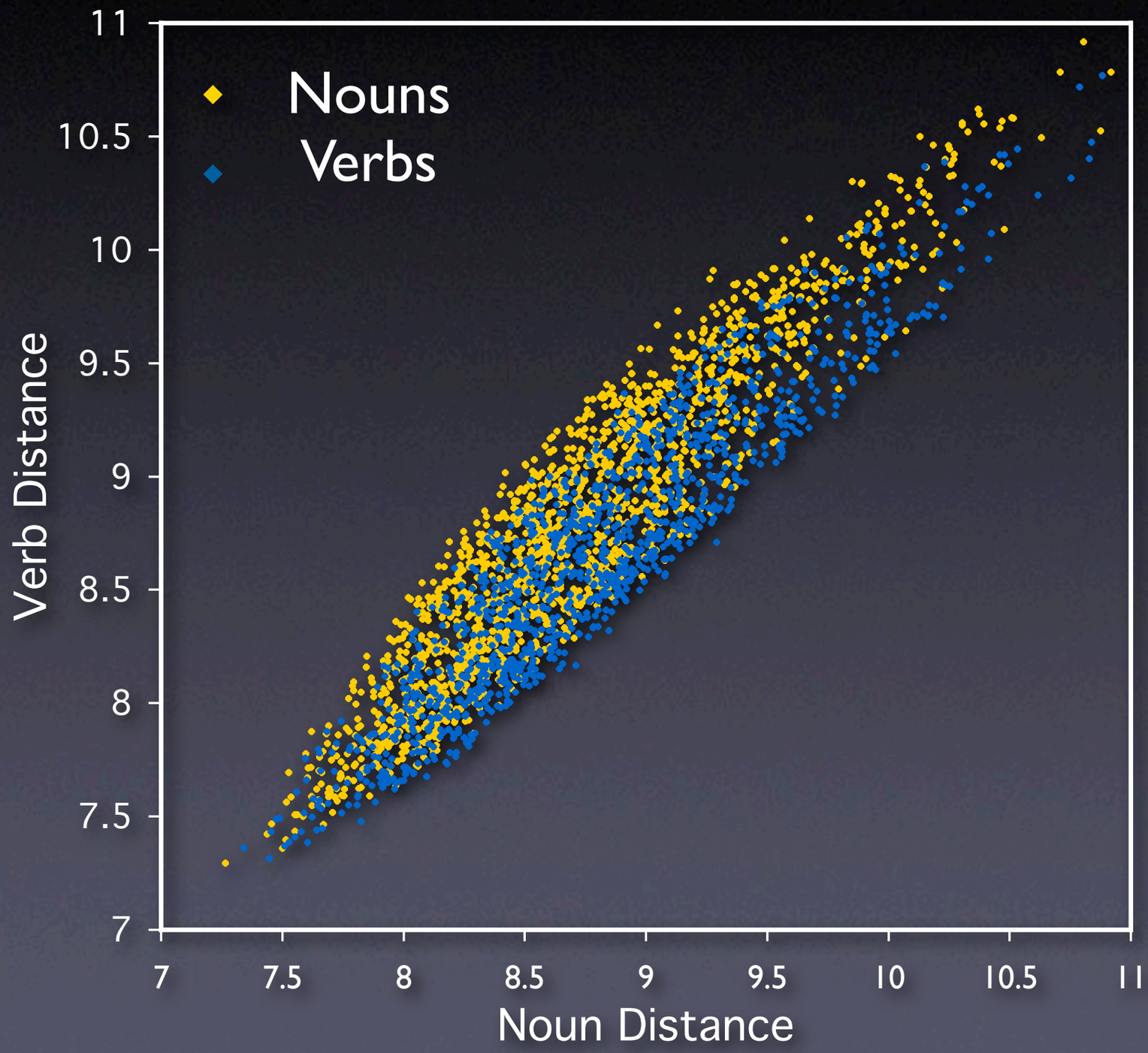
- Whale: "oblong viviparous fish"
- Dog: "rapacious viviparous animal with an oblong head"

John Wilkins (1614-1672)

# Sound Symbolism Today

- Speech gesture cues to meaning (Howard)
- Prosodic cues to word meaning (Laura)
- Prosodic cues to social coordination (Sandy)
- Phonaesthetic cues to meaning (Ben)
- Cross-linguistic phonological cues to meaning (Lynne)
- Mimetic cues to meaning (Mustumi)
- Synaesthetic grounding of sound symbolic cues (Daphne)
- Perceptual cues to metaphor (Krish)

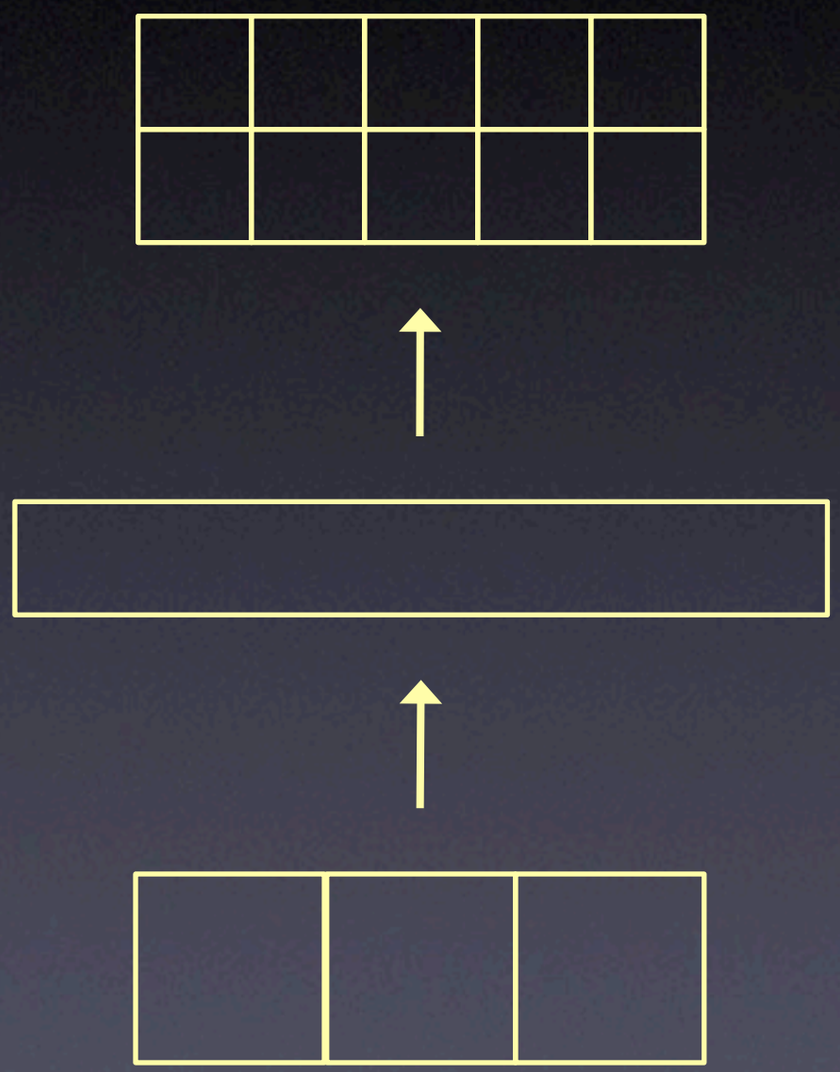
# Phonological Cues to Lexical Categories




Source:
Farmer,
Christiansen &
Monaghan,
*PNAS*, 2006

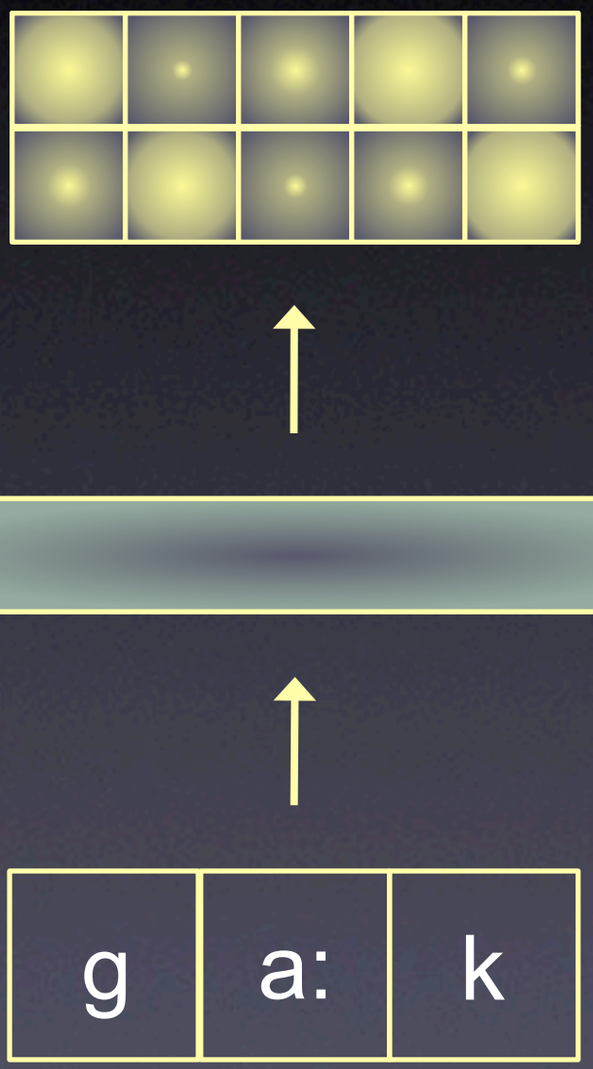
# Can systematic sound-meaning mapping facilitate language acquisition?

# Systematicity and Word Learning



- Two sets of input words
  - Plosives/back vowels:
    - *ga:k, gu:k, ka:g, ku:g, ka:k, gu:g*
  - Fricatives/front vowels:
    - *fi3, fl3, 3if, 3if, 3i3, flf*
- Two categories of output
  - 0.25 activation prototype
  - 0.75 activation prototype

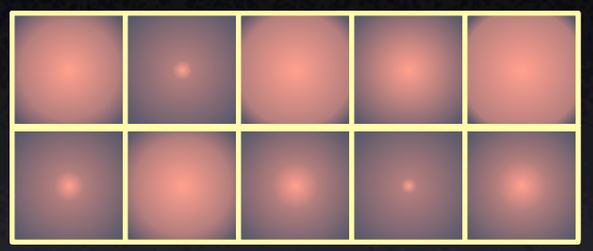
# Systematic Mapping




0.75 activation prototype

plosives/back vowels

# Systematic Mapping



f | i | 3

0.25 activation prototype

fricatives/front vowels

0.75 activation prototype

plosives/back vowels

# Arbitrary Mapping

0.75 activation prototype

0.25 activation prototype

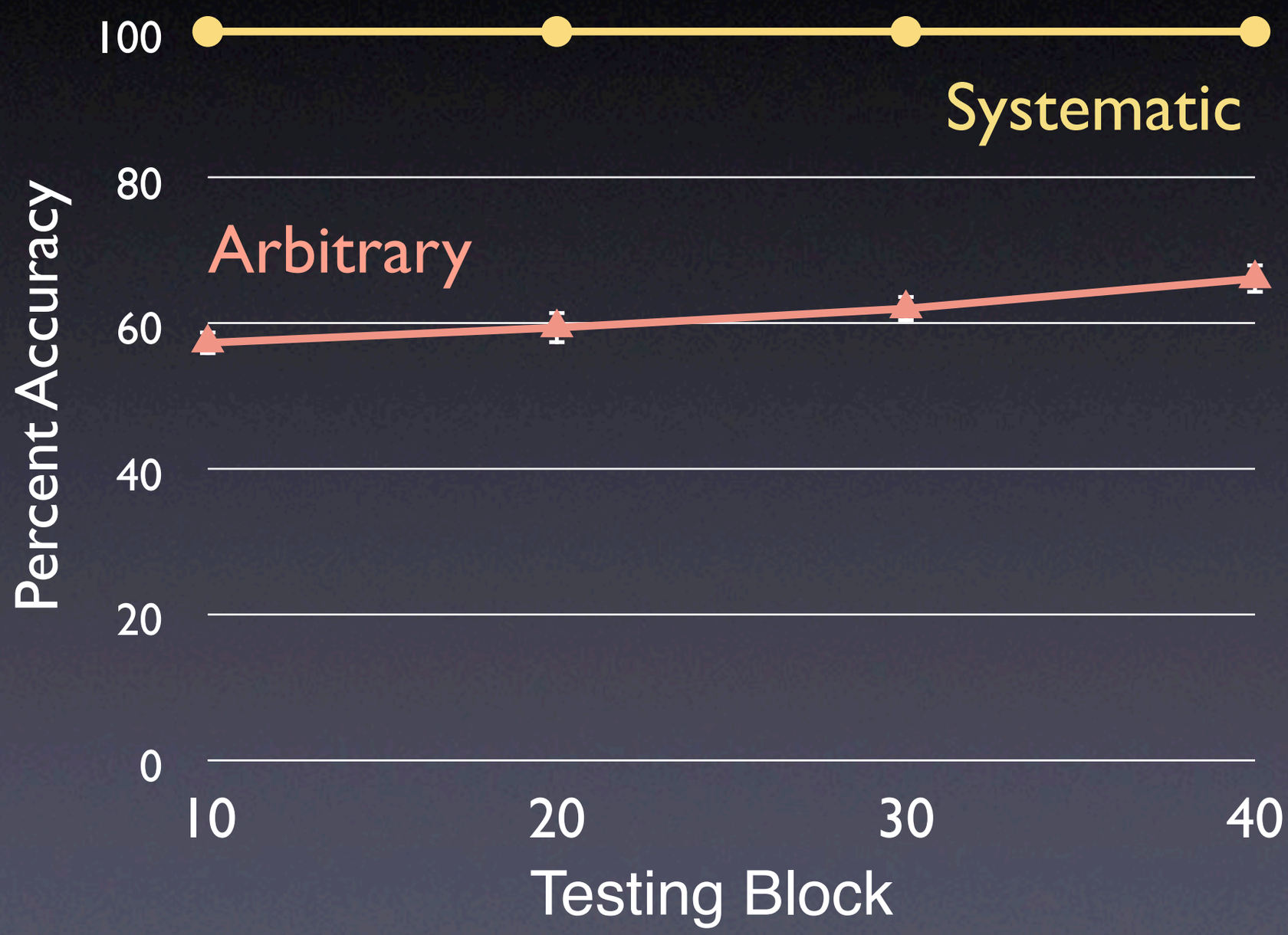
fricatives/front vowels

0.25 activation prototype

0.75 activation prototype

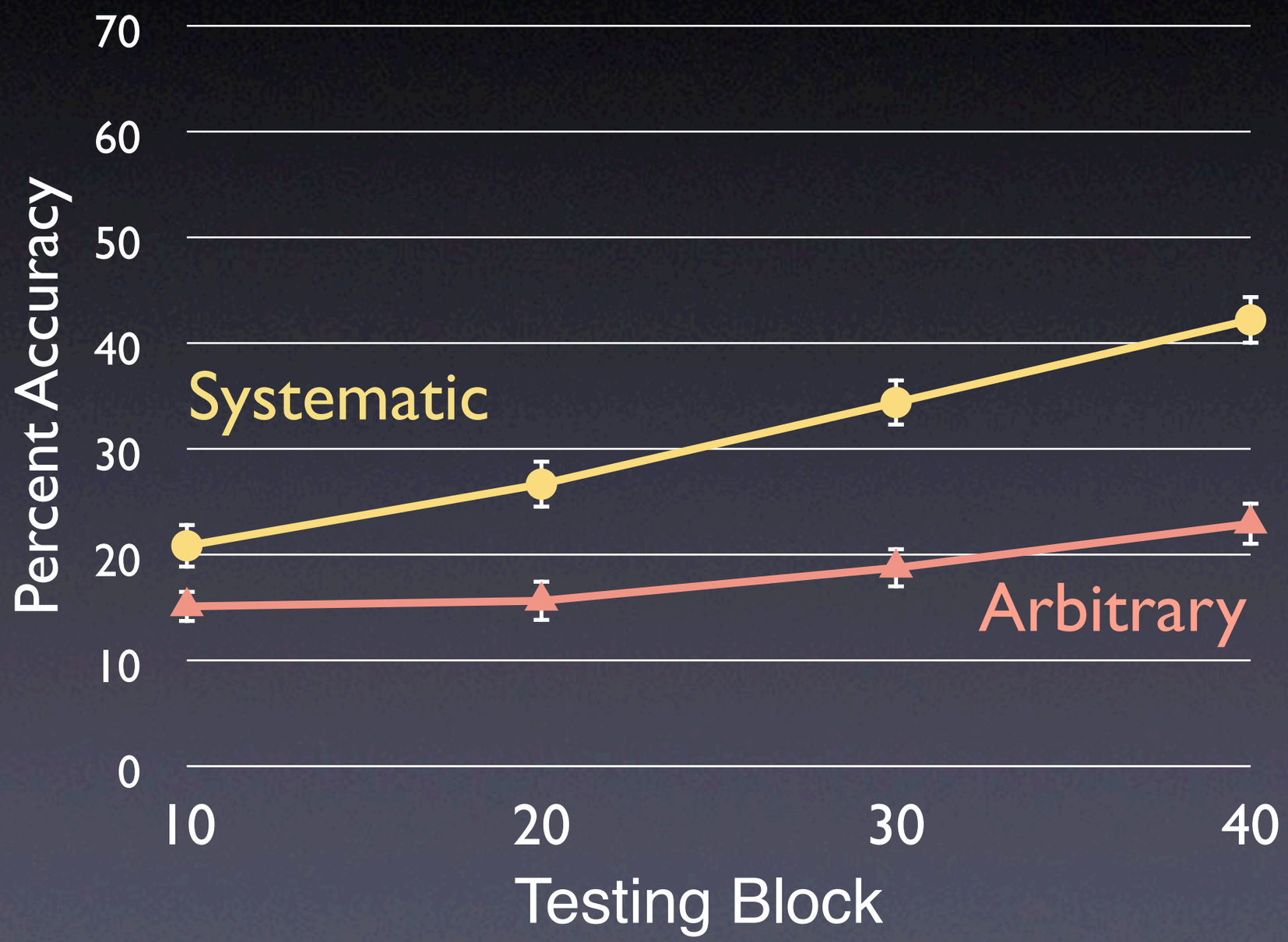
plosives/back vowels

# The Systematic Advantage
# Category Learning




Source: Monaghan, Christiansen & Fitneva, in preparation

# The Systematic Advantage
# Word Individuation




Source: Monaghan, Christiansen & Fitneva, in preparation

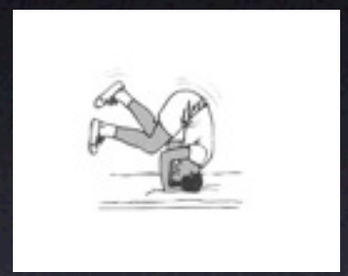
Do human learners show a systematic advantage?

# Systematic Condition

*ga:k*

*fiʒ*

*gu:k*



*fIʒ*

.
.
.

*ka:g*

.
.
.

*ʒif*

*gu:g*

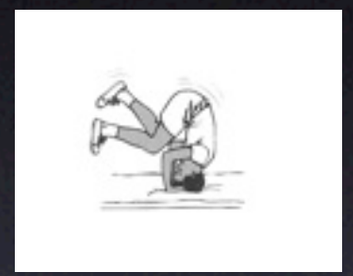
*fIf*

# Arbitrary Condition

*ga:k*

*gu:k*

*fi3*



*fI3*

.
.
.

*3if*

*fIf*

.
.
.

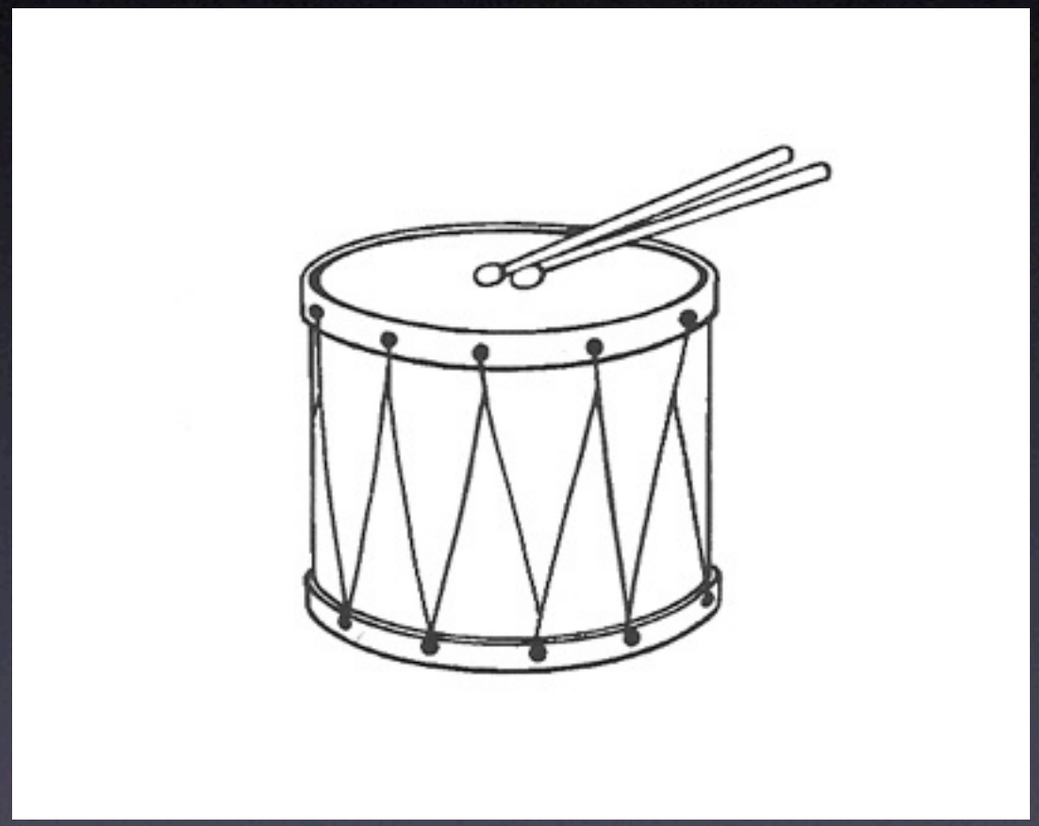
*ka:g*

*gu:g*

# Actions

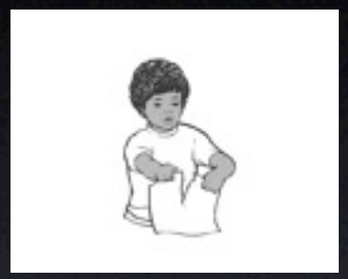
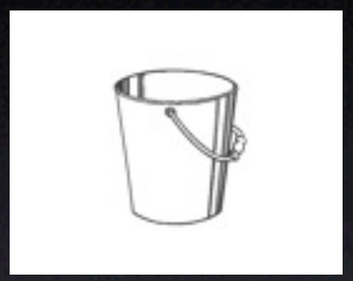
*ga:k*

# Nouns

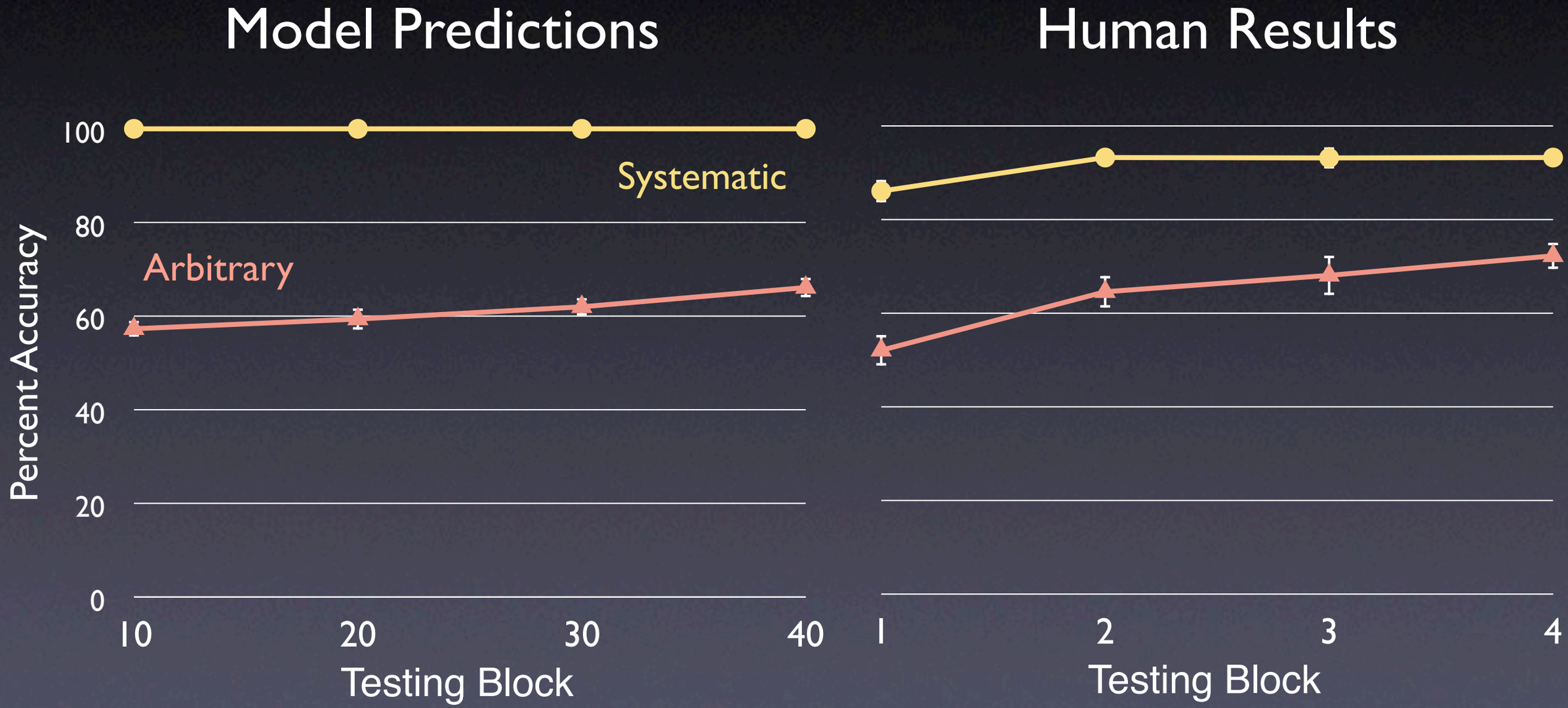


*fIf*

# Test Procedure

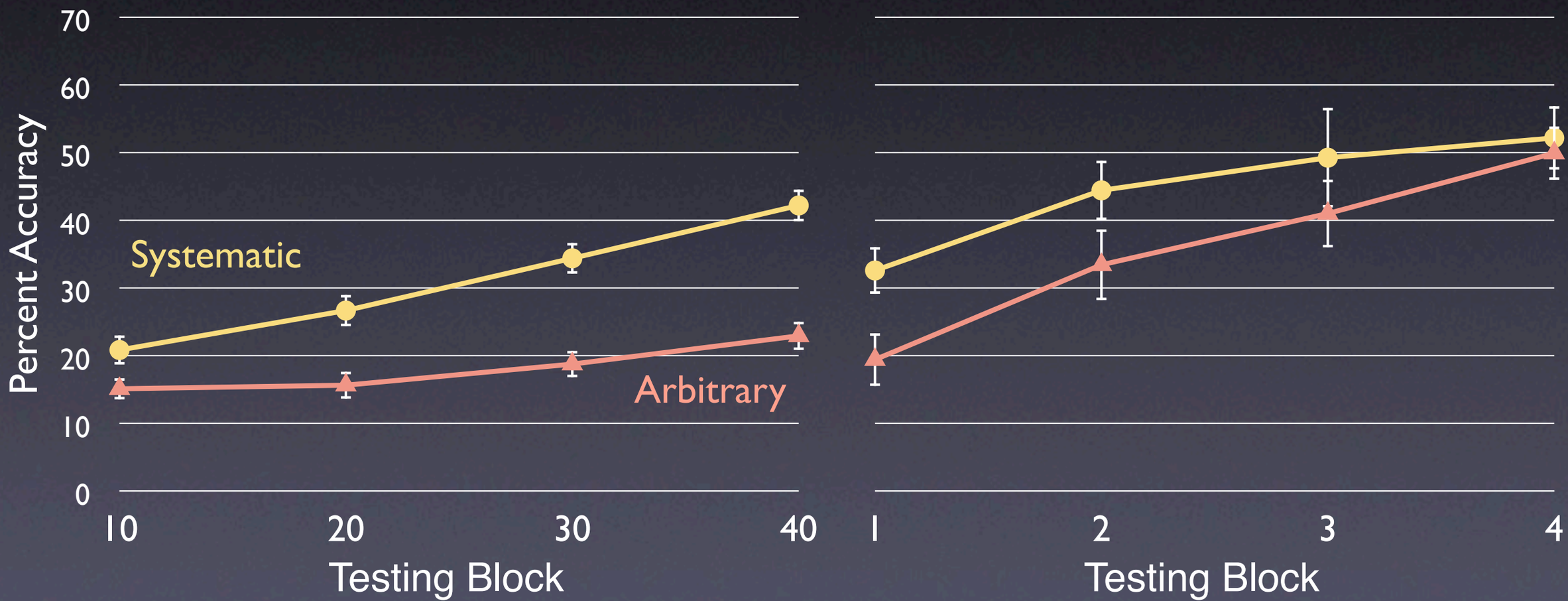



# The Systematic Advantage
# Category Learning




Source: Monaghan, Christiansen & Fitneva, in preparation

# The Systematic Advantage
# Word Individuation

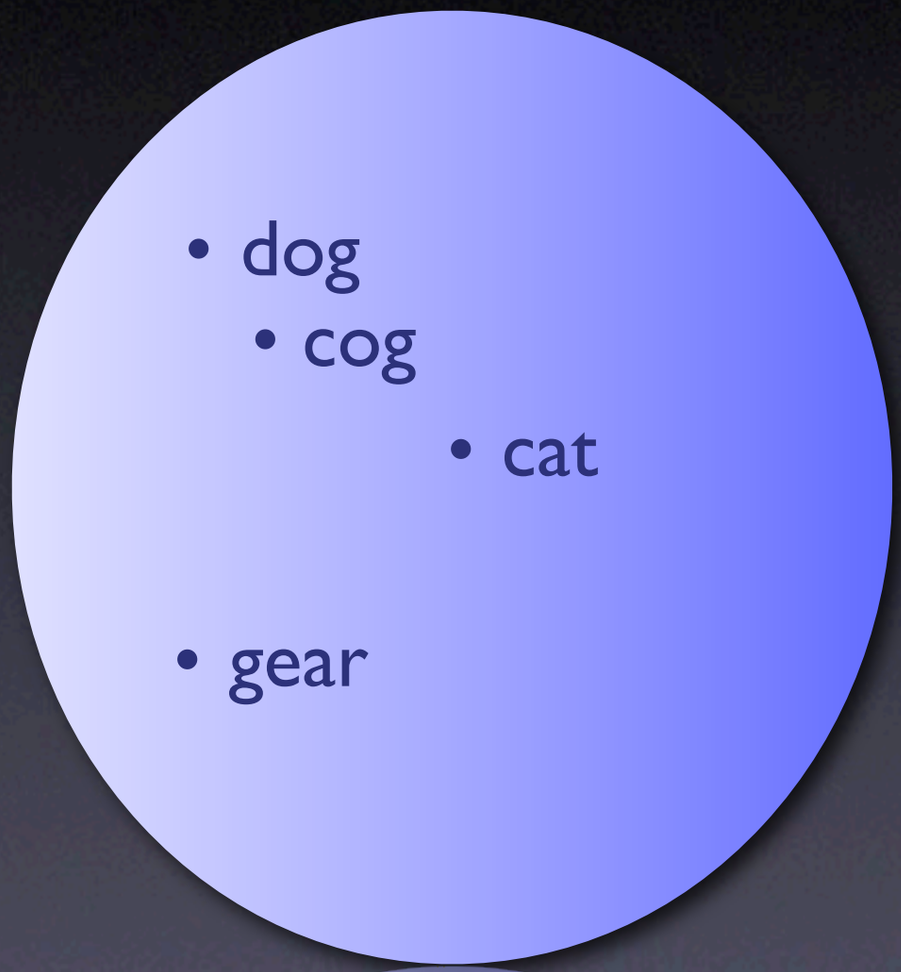



Source: Monaghan, Christiansen & Fitneva, in preparation

# How prevalent are systematic form-meaning mappings?

# How Systematic Is Language?

6028 English monosyllabic words




Phonological space

[phonological features from CELEX]

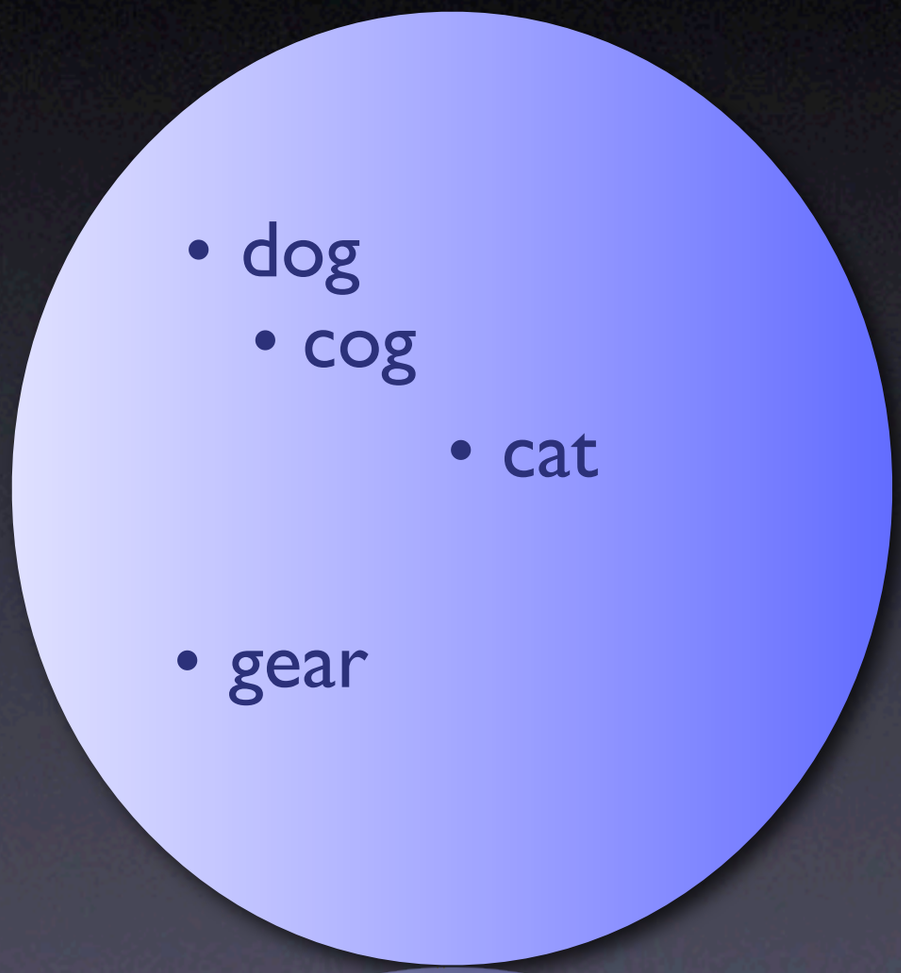
Semantic space

[contextual co-occurrence vectors]

# Testing Systematicity

- 6028 x 6028 matrix of similarities
- Correlate phonological similarities and meaning similarities
  - $r = 0.05$, $p < .0000000000\ldots001$
- Mantel test: swap two rows/columns and see if correlation reduces or increases
  - $p = .338$

# Could Language Be More Systematic or More Arbitrary?




Phonological space

Semantic space

# Manipulating Arbitrariness

- More systematic?
  - swap two rows/columns, if correlation increases keep the swap
    - After 500,000 swaps: $r$ = .495
- More arbitrary?
  - Swap two rows/columns, if correlation decreases keep the swap
    - After 120,000 swaps: $r$ = 0

# Why isn't sound symbolism more prevalent in language?

# Problems with Systematicity?

- Easy to mistake similar sounding words

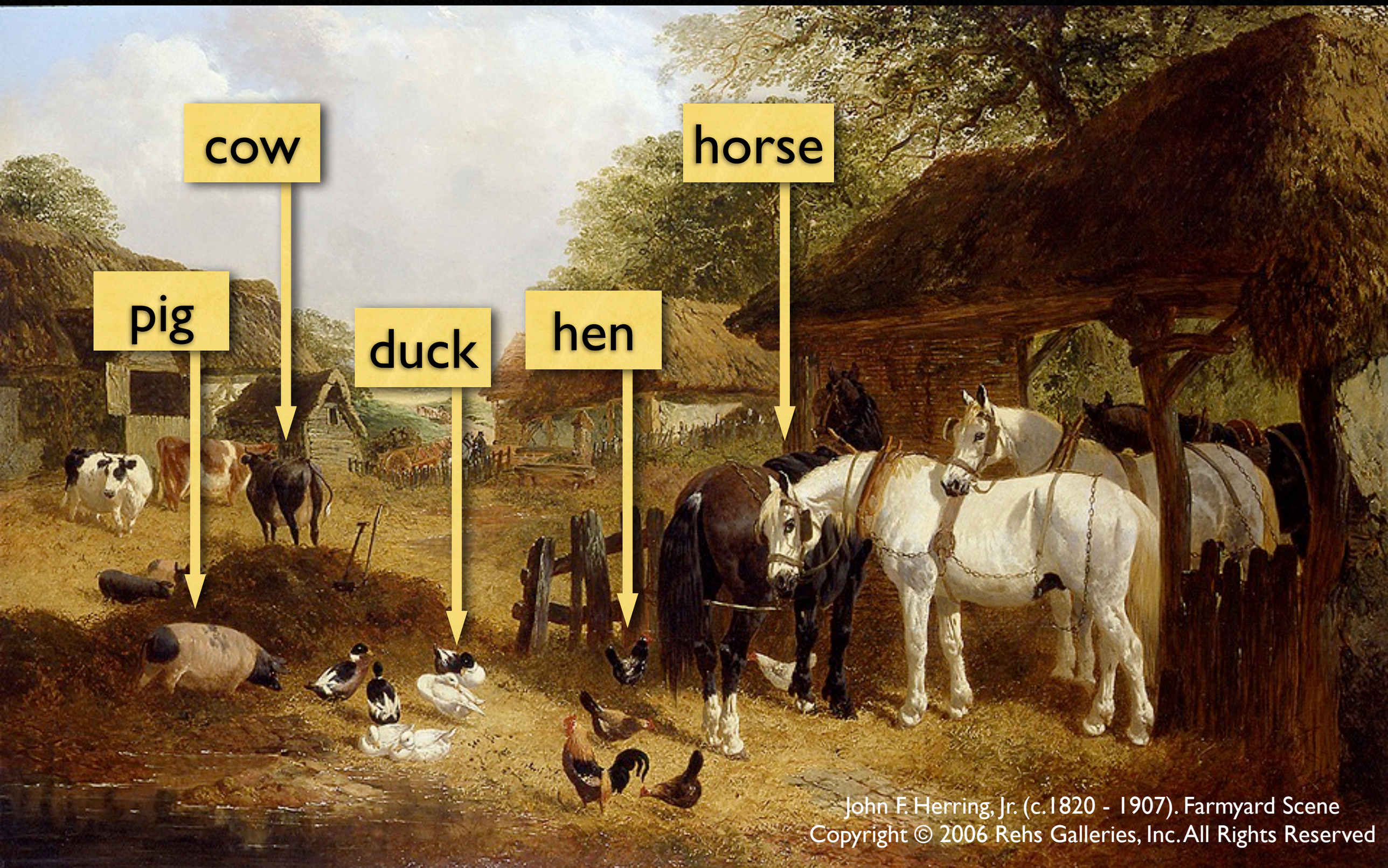
Wilkins miswrites "gade" (barley), for "gape" (tulip)

Eco (1995)

"edible plants or berries could be confused with poisonous ones, and animals that attack could be confused with those that are benign"

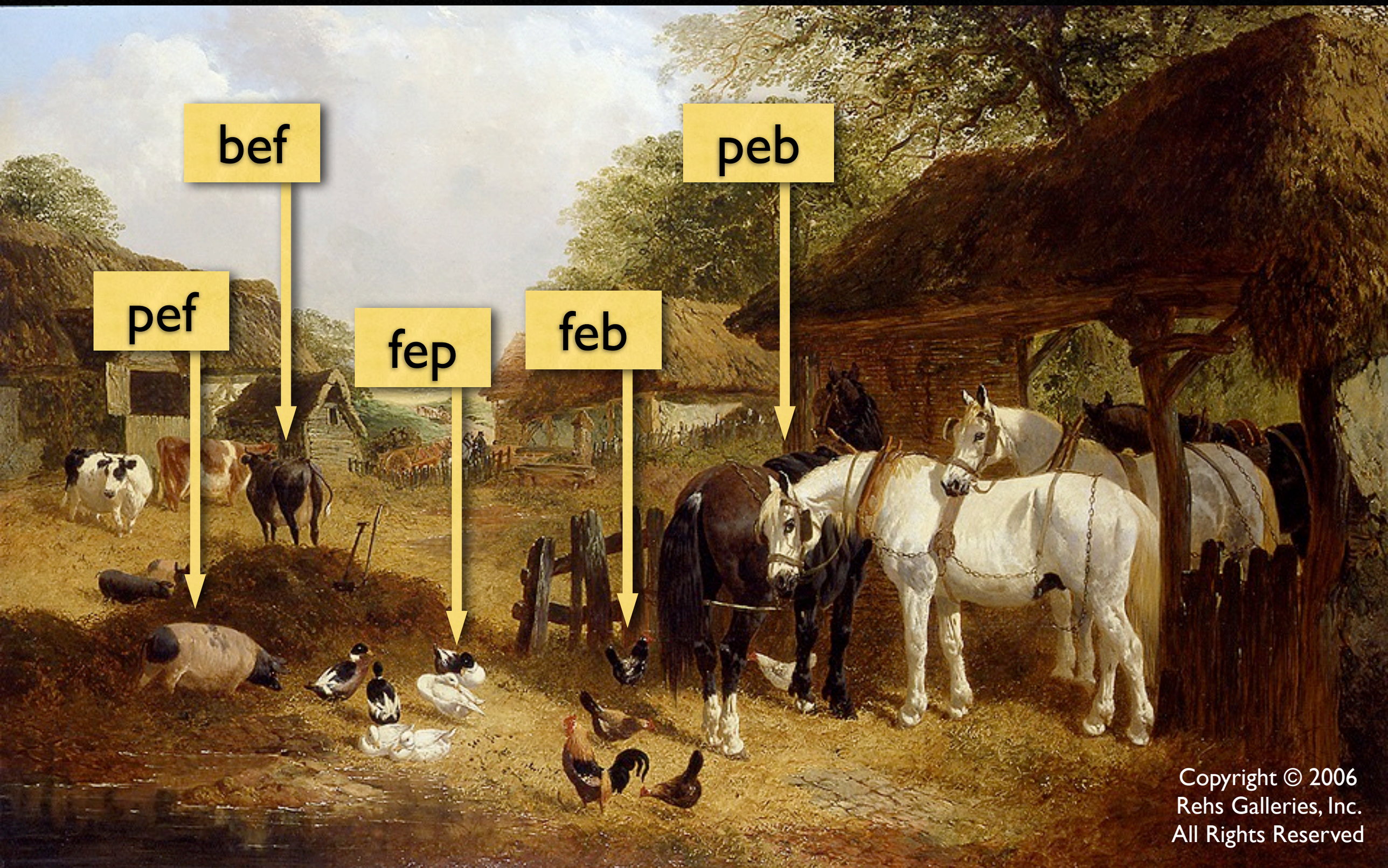
Corballis (2002: p.187)

# Arbitrary Learning Advantage?




John F. Herring, Jr. (c.1820 - 1907). Farmyard Scene
Copyright © 2006 Rehs Galleries, Inc. All Rights Reserved

# Systematic Learning Disadvantage?




Copyright © 2006
Rehs Galleries, Inc.
All Rights Reserved

# Advantages of Arbitrariness?

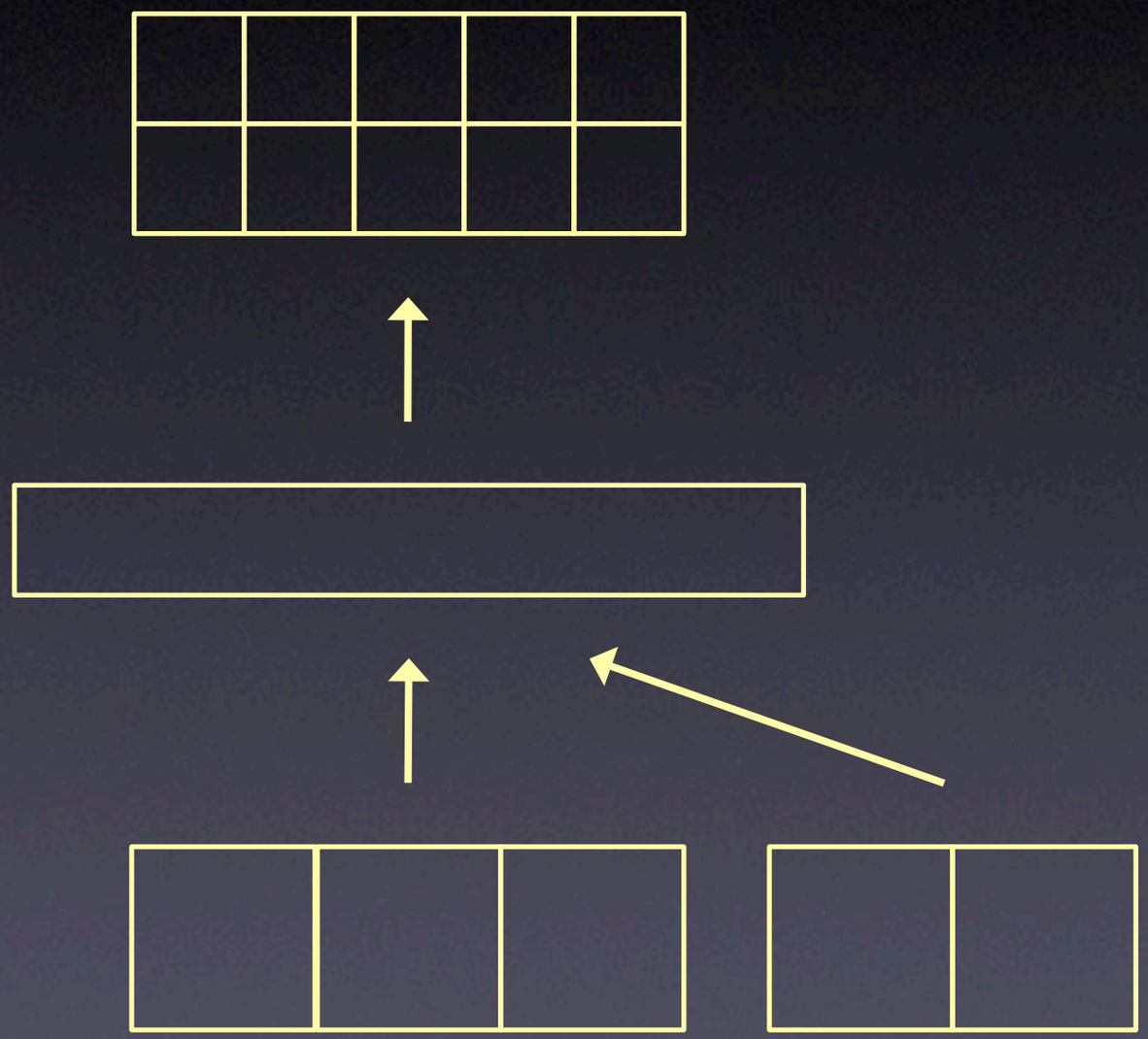
- Maximizes degrees of freedom
  - Learning advantage for large vocabularies
    (Gasser, 2004)
- Maximizes distinctiveness
  - Context gives general idea of meaning
  - Phonology identifies precise referent, precise meaning
- Minimizes effects of noise in the input

# Working Hypotheses:

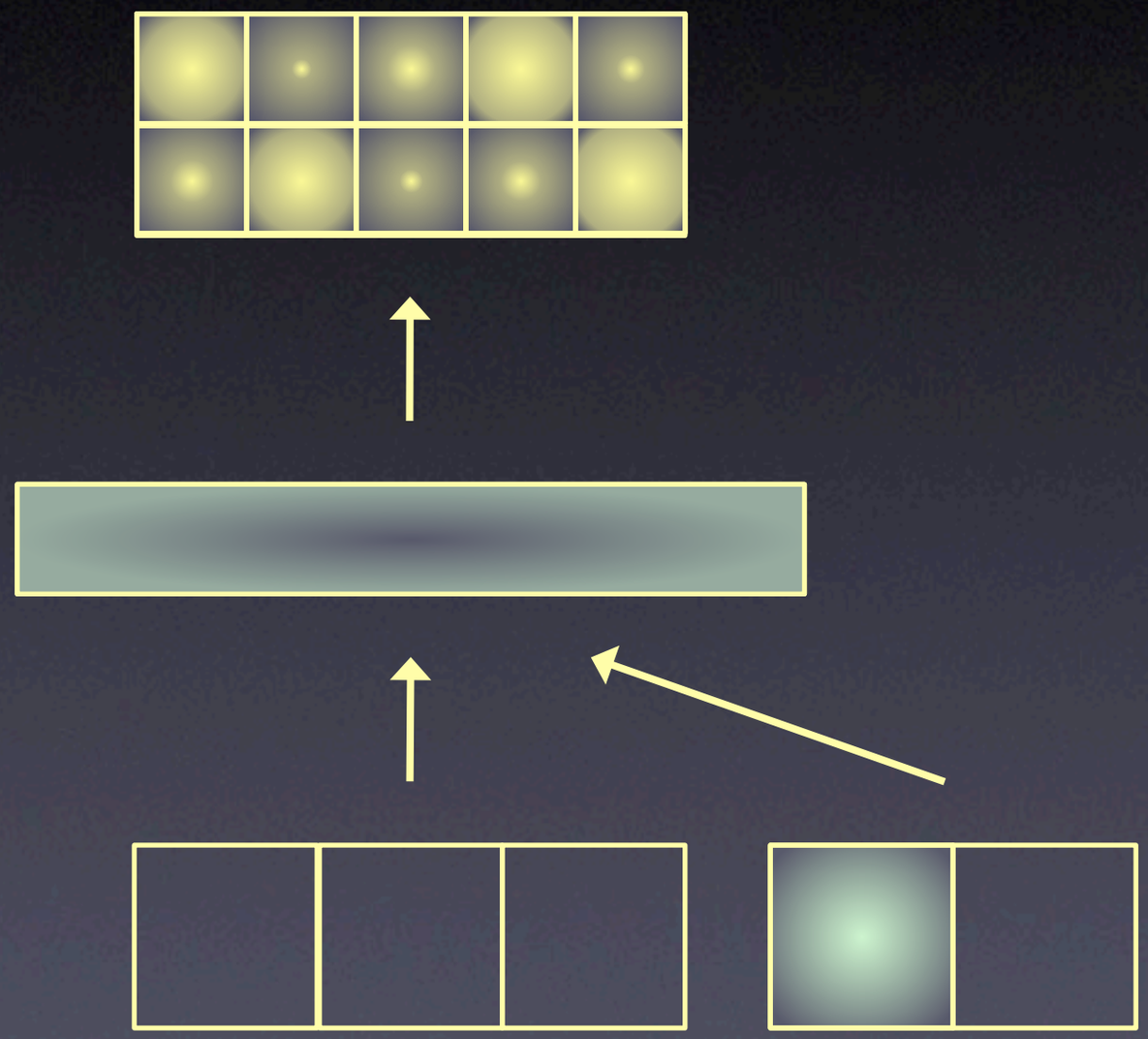
- Words are not learned in isolation but as part of a language system
  - Learning the meaning of words and how to use them are not separate tasks
  - A constellation of cues to the referent is available to the listener
  - Arbitrariness maximizes the distinctiveness of the sound of a word and this facilitates perception, and therefore also learning

Might context bring out an arbitrary advantage?

# Simulating Word Learning in Context

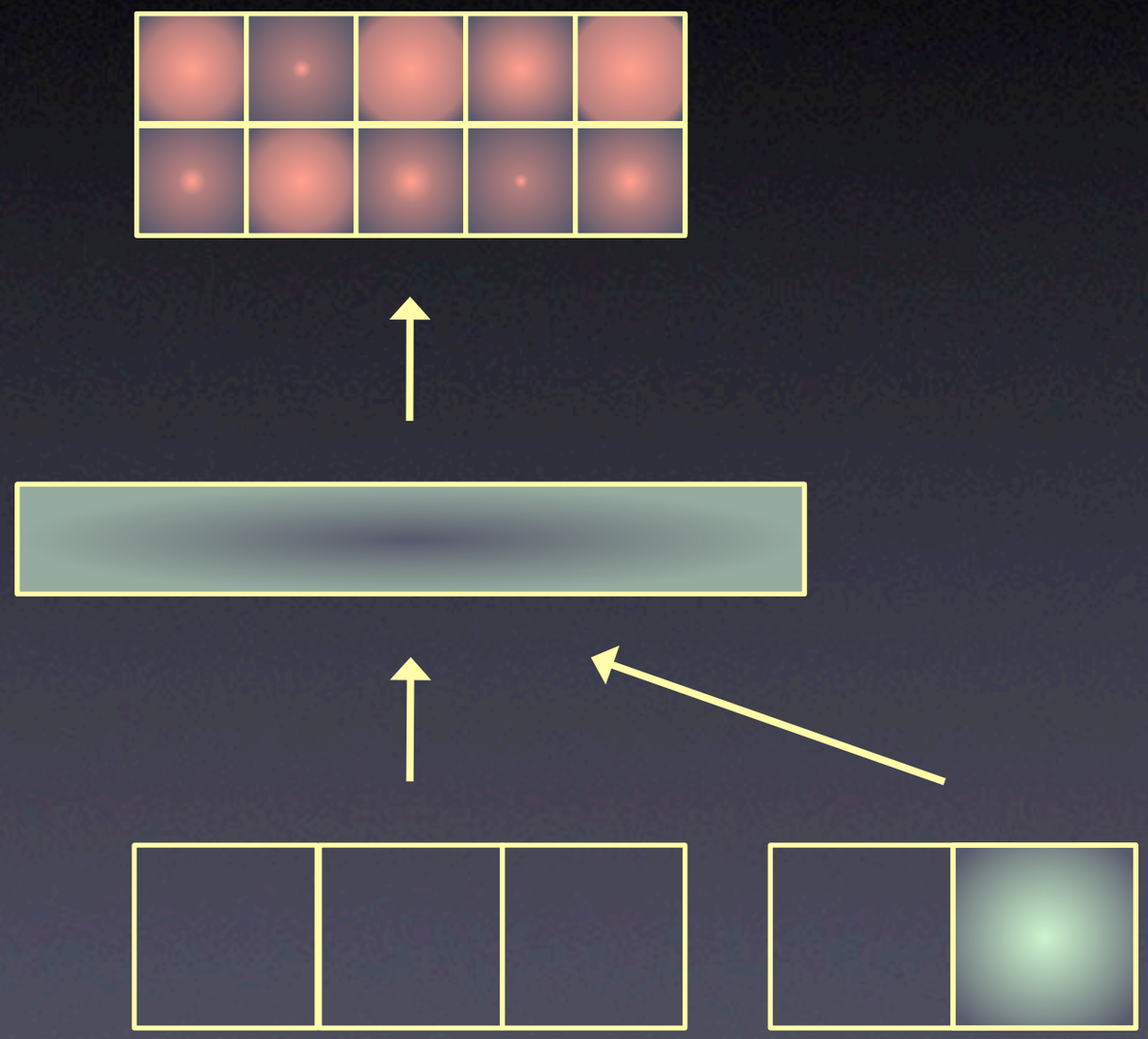


# Simulating Word Learning in Context



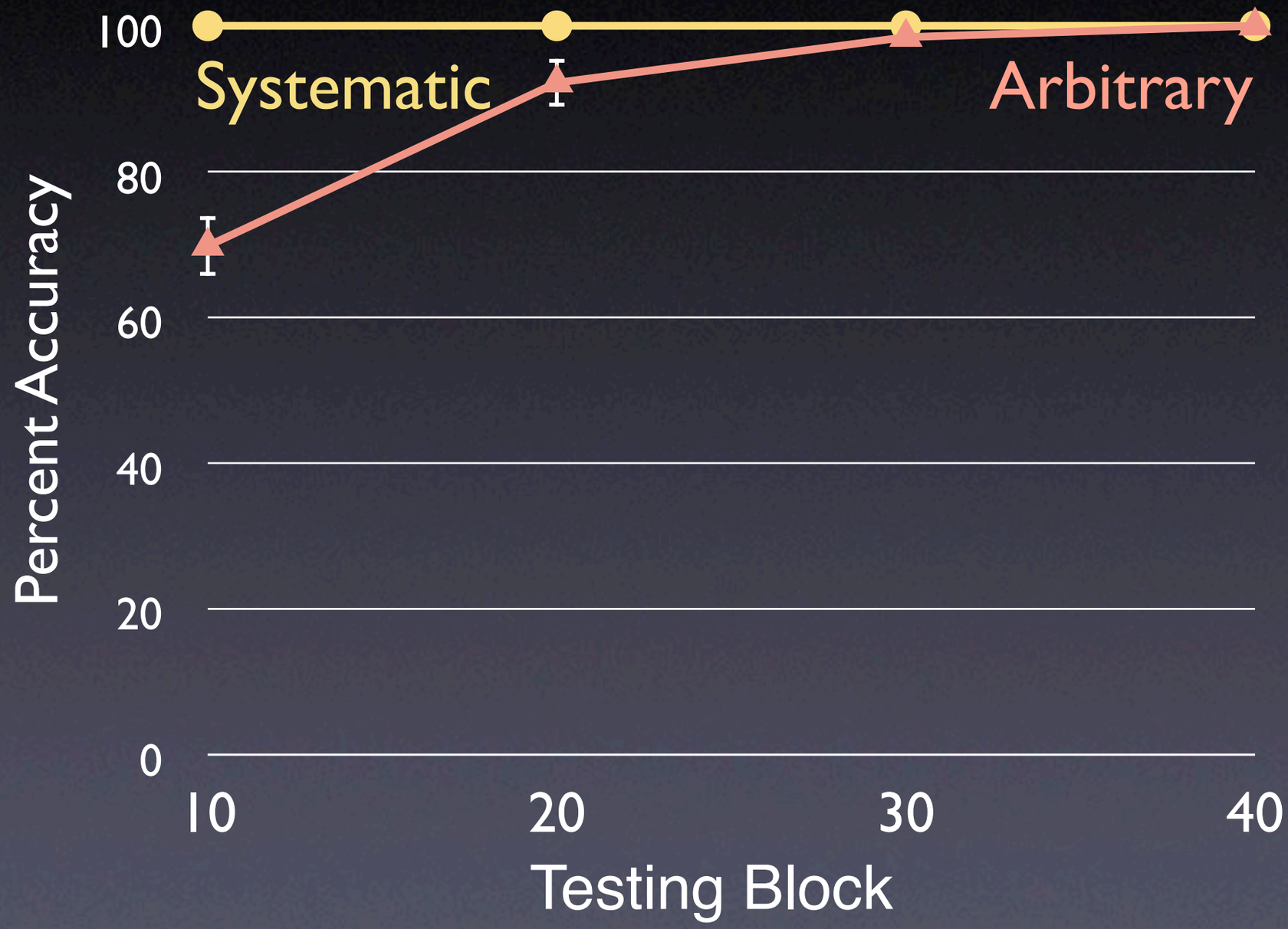
0.75 activation prototype

# Simulating Word Learning in Context



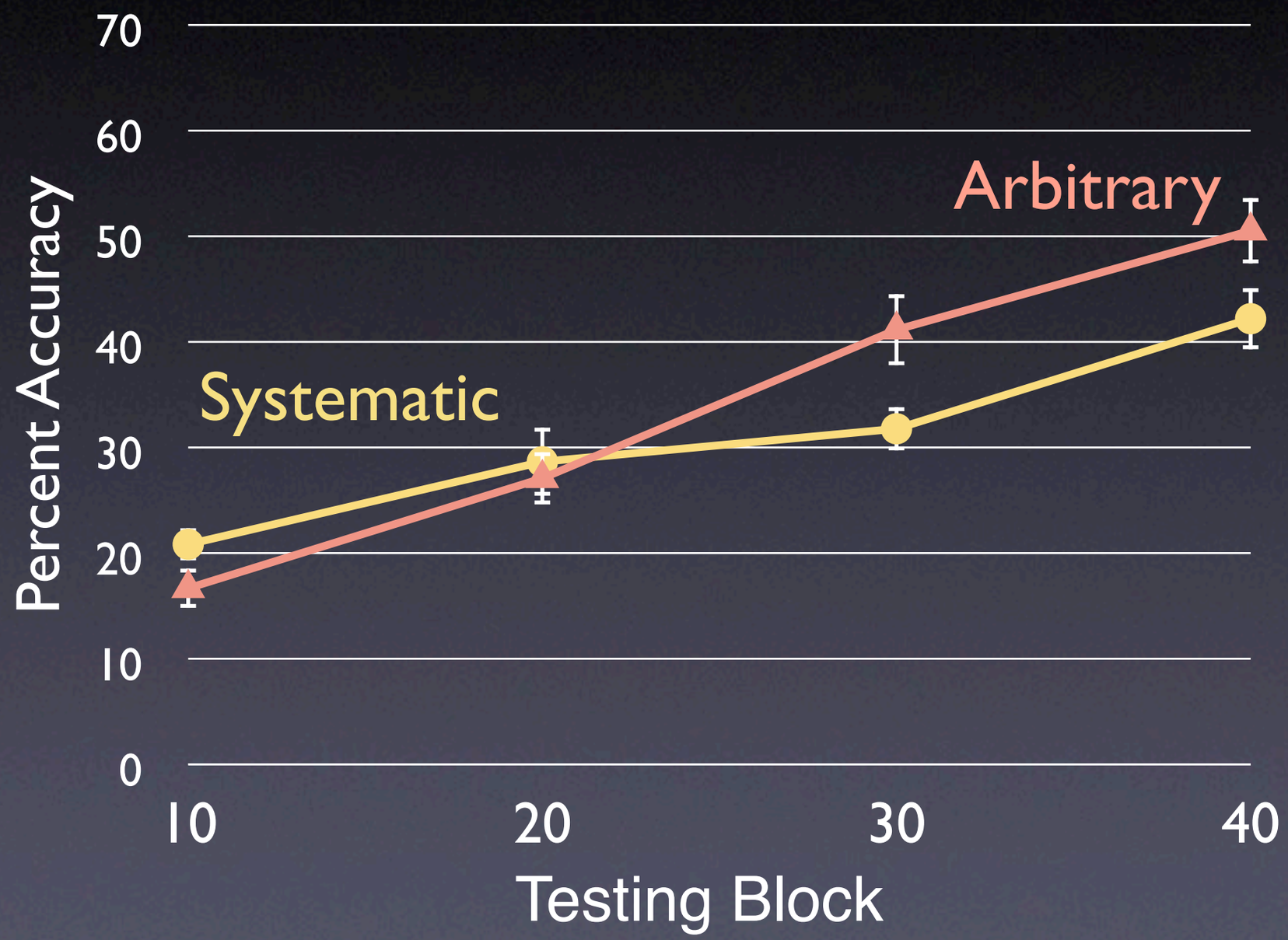
0.25 activation prototype

# The Arbitrary Advantage
# Category Learning with Context




Source: Monaghan, Christiansen & Fitneva, in preparation

# The Arbitrary Advantage
# Word Individuation with Context




Source: Monaghan, Christiansen & Fitneva, in preparation

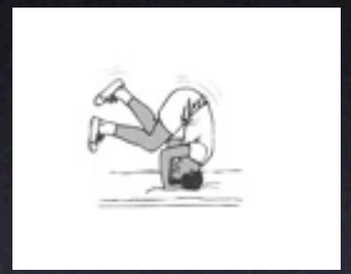
Do human learners show an arbitrary advantage?

# Systematic Condition

wɛ *ga:k*

wɛ *gu:k*

*mə fiʒ*



*mə fIʒ*

.
.
.

wɛ *ka:g*

wɛ *gu:g*

.
.
.

*mə ʒif*

*mə fIf*

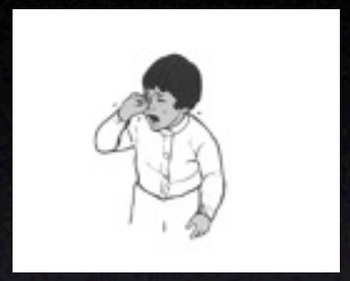
# Arbitrary Condition

wɛ *ga:k*

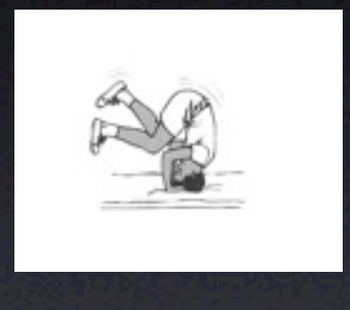
wɛ *gu:k*

wɛ ʒ*iʃ*

wɛ *ʃIʃ*



*mə ʃiʒ*



*mə ʃIʒ*

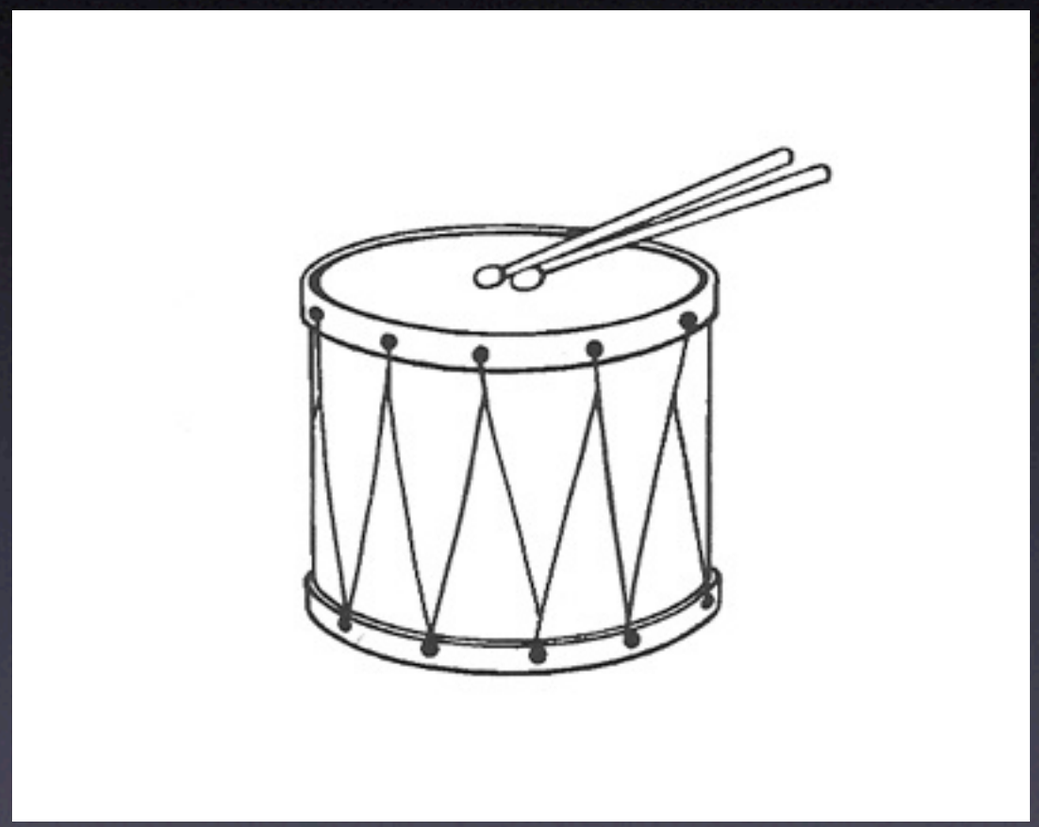
*mə ka:g*

*mə gu:g*

# Actions

*weh ga:k*

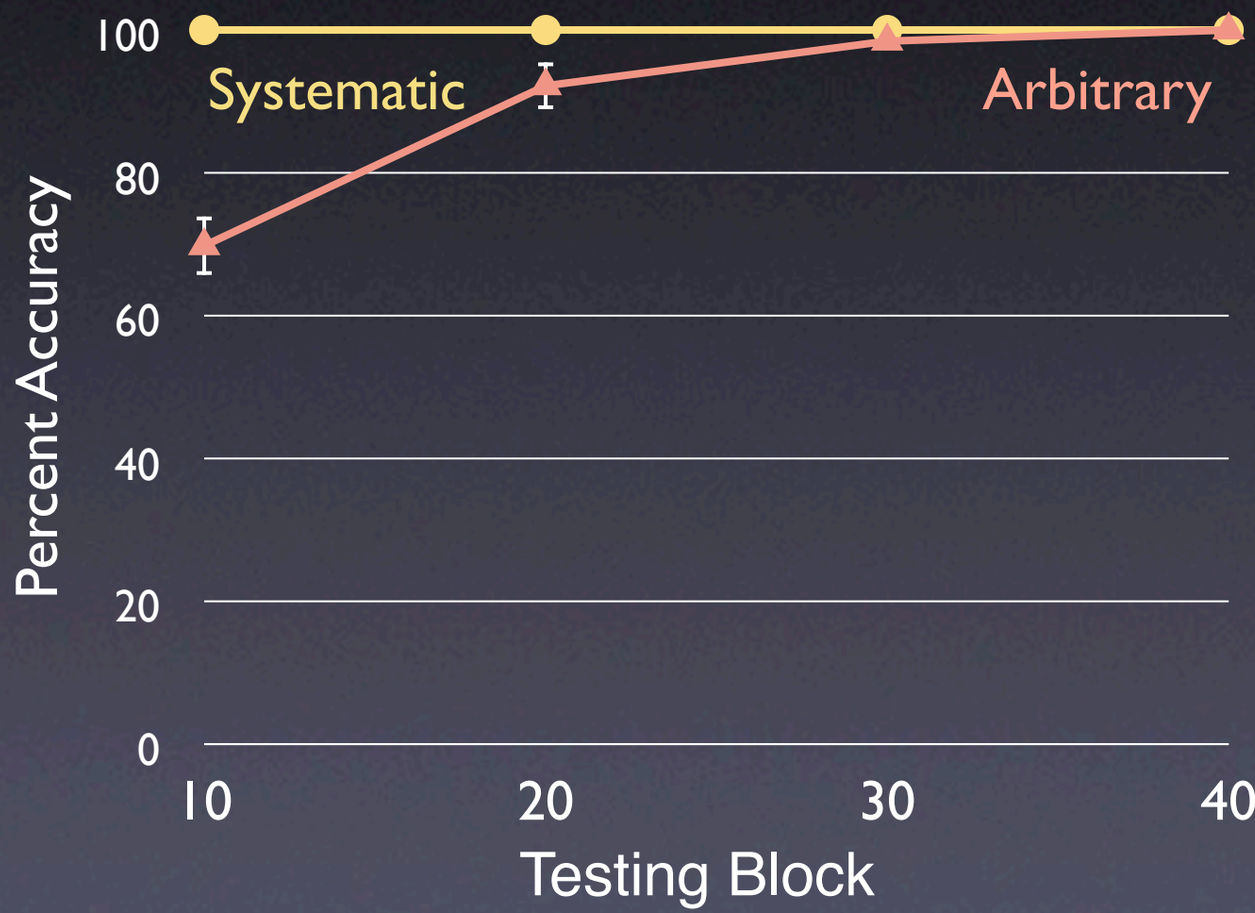
# Nouns



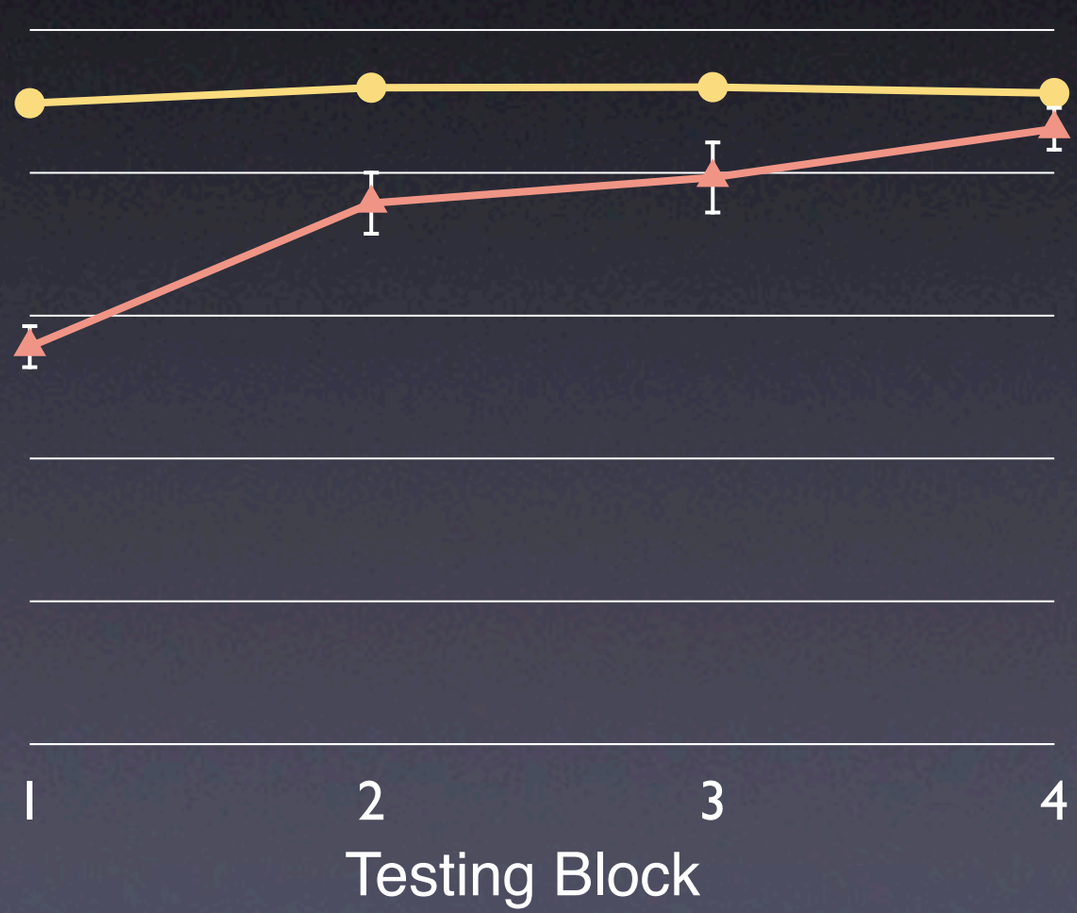
*muh fIf*

# The Arbitrary Advantage
# Category Learning with Context

## Model Predictions



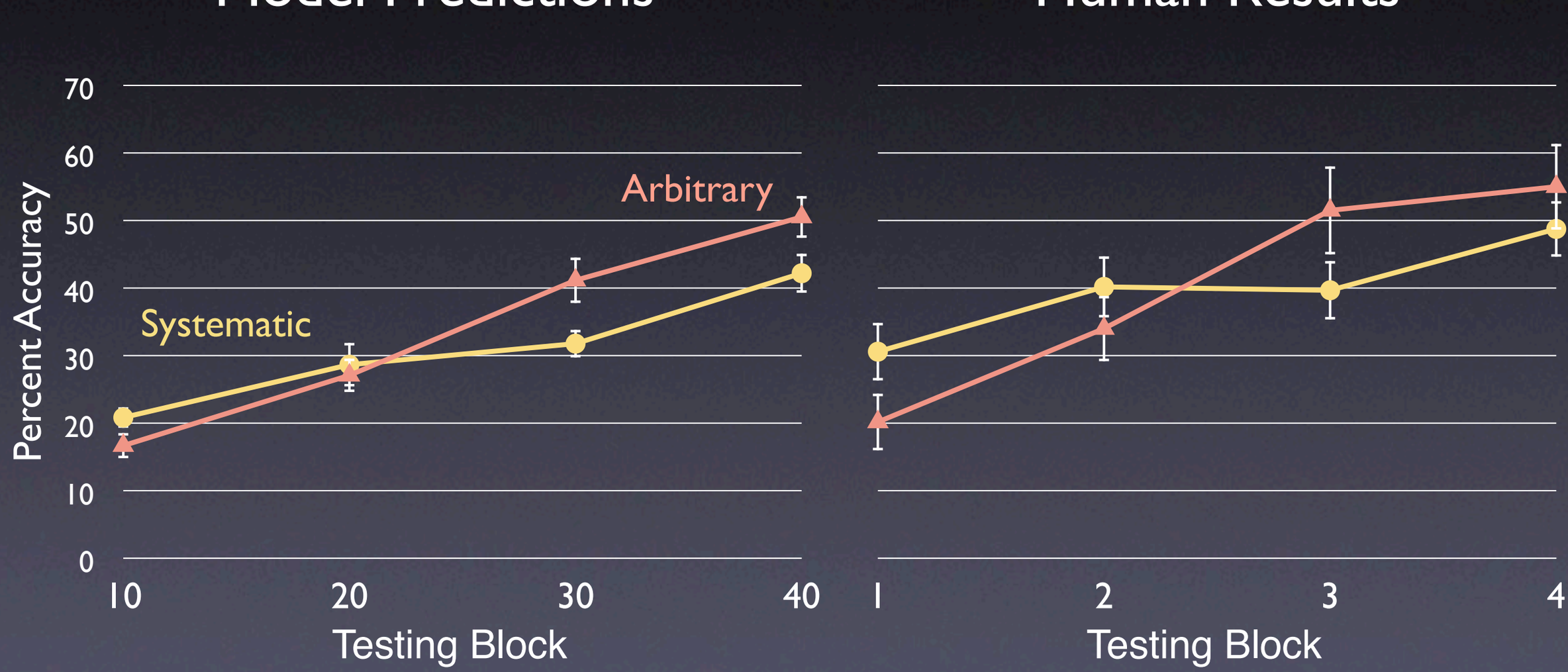

## Human Results




Source: Monaghan, Christiansen & Fitneva, in preparation

# The Arbitrary Advantage
# Word Individuation with Context






Source: Monaghan, Christiansen & Fitneva, in preparation

# Division of Labor

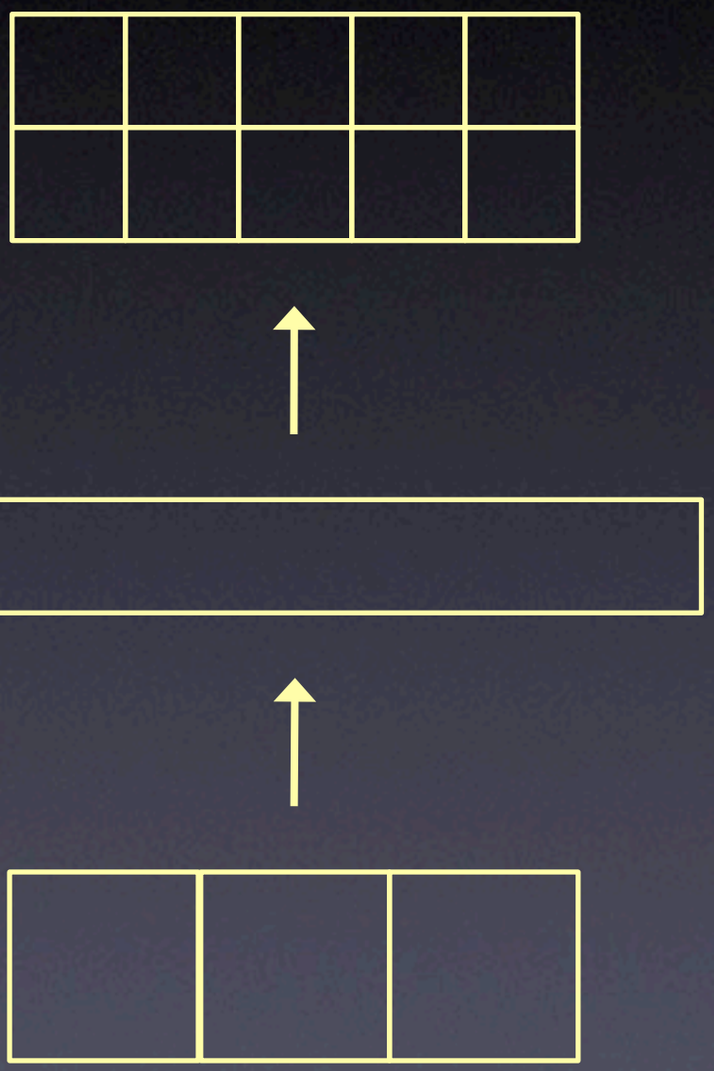
- Systematicity maximizes recycling of previous information
- Arbitrary mappings facilitate meaning individuation in context

# Division of Labor within Words?

- Distributional context can facilitate the learning of arbitrary mappings
- But distributional context is not reliable for novel and low-frequency words
  (Monaghan, Chater & Christiansen, 2005)
- Perhaps both arbitrary and systematic components co-exist within a word?

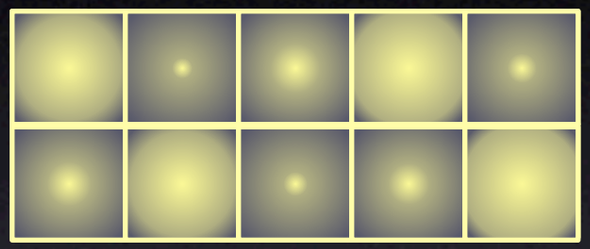
Might an arbitrary-systematic division of labor within words work?

# A Division of Labor Model



- Arbitrary information at the beginning of words
- Systematic information at the end
  - Plosives as codas:
    - *ga:k, gu:g, ka:g, ʒik, ʒig, flk*
  - Fricatives as codas:
    - *ku:f, ka:ʒ, gu:f, fiʒ, flʒ, ʒif*

# Half-Half Mapping

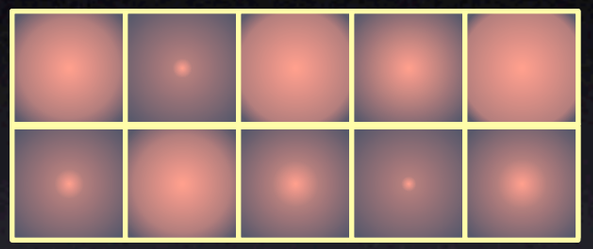


g | a: | k

0.75 activation prototype

plosive codas

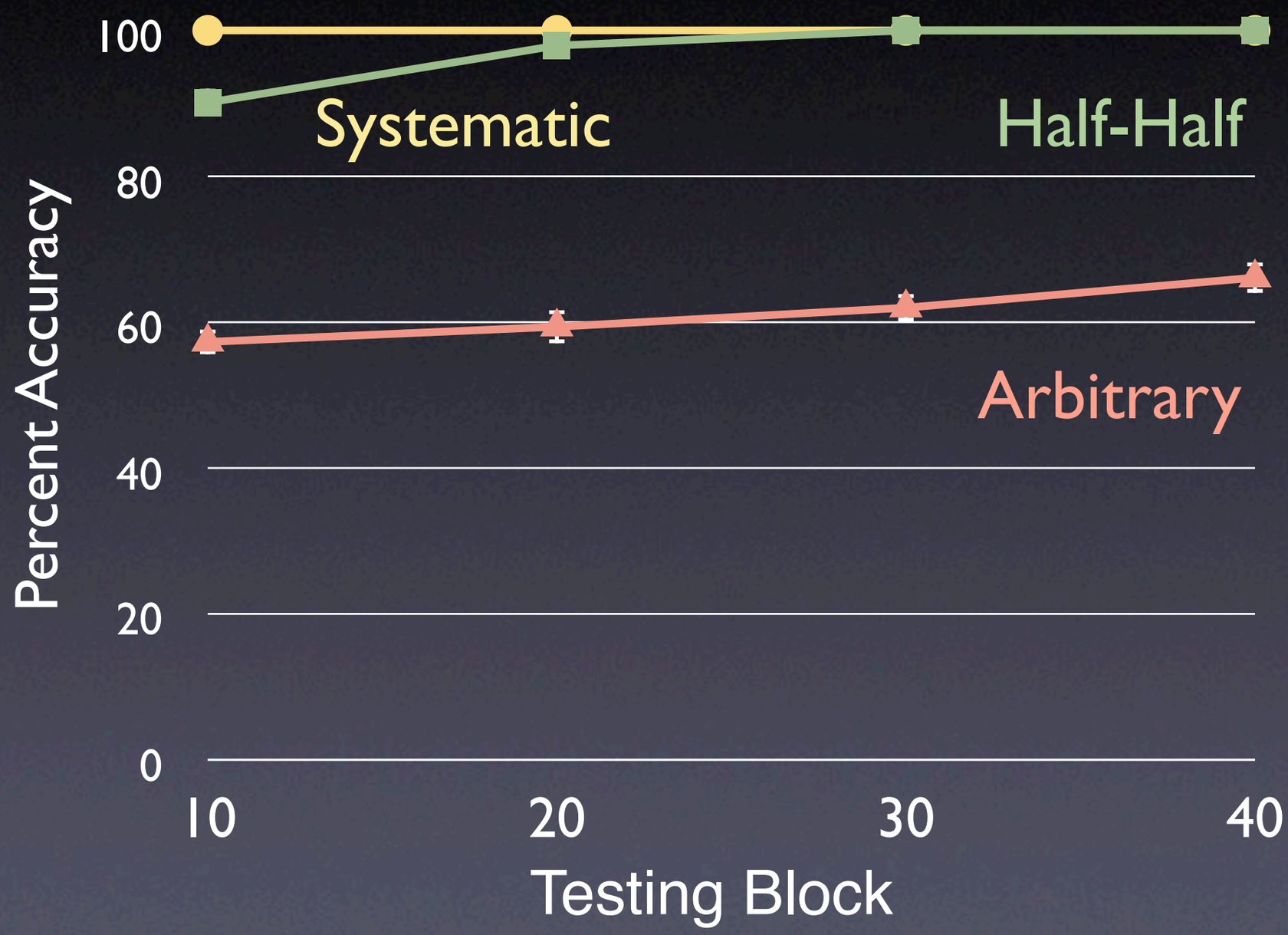
# Half-Half Mapping



0.25 activation prototype

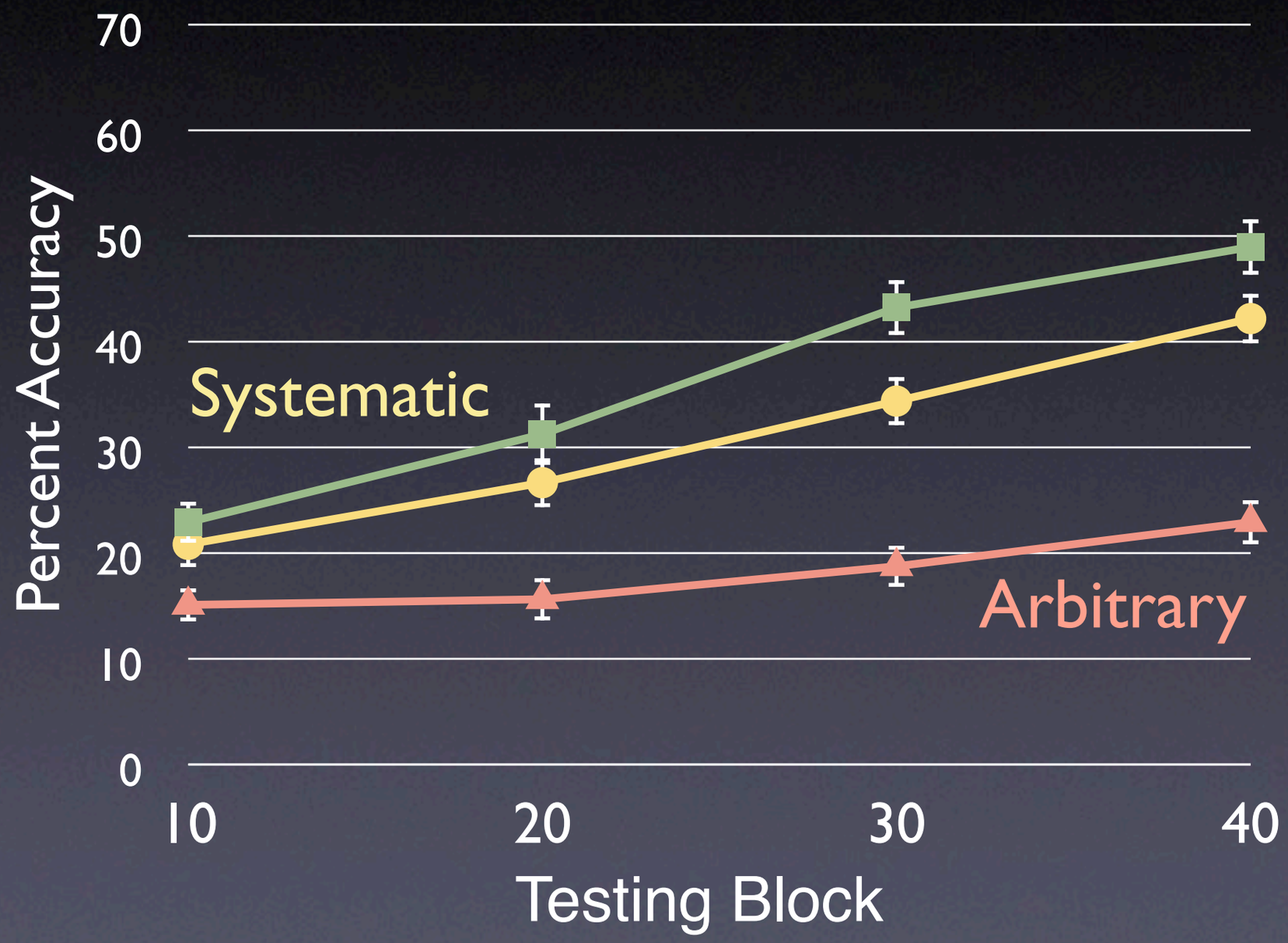
fricative codas

# Division of Labor
# Category Learning




Source: Monaghan, Christiansen & Fitneva, in preparation

# Division of Labor
# Word Individuation



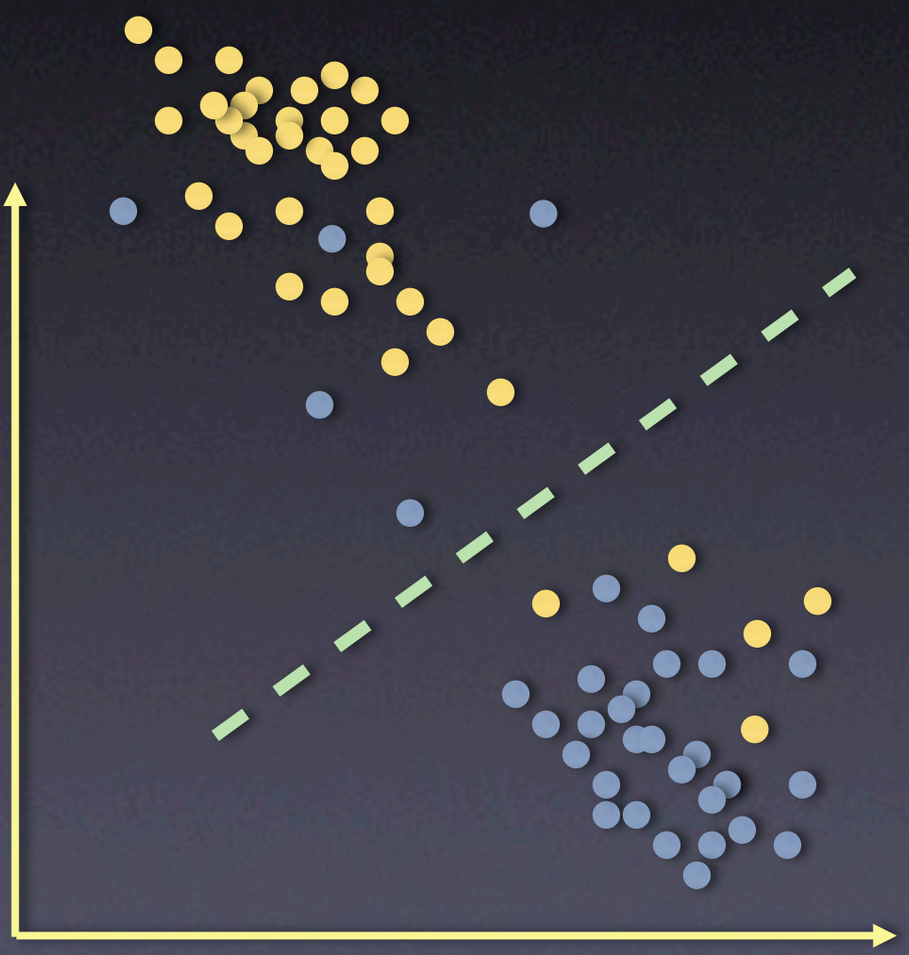

Source: Monaghan, Christiansen & Fitneva, in preparation

Do real words incorporate a division of labor?

# A Corpus Analysis

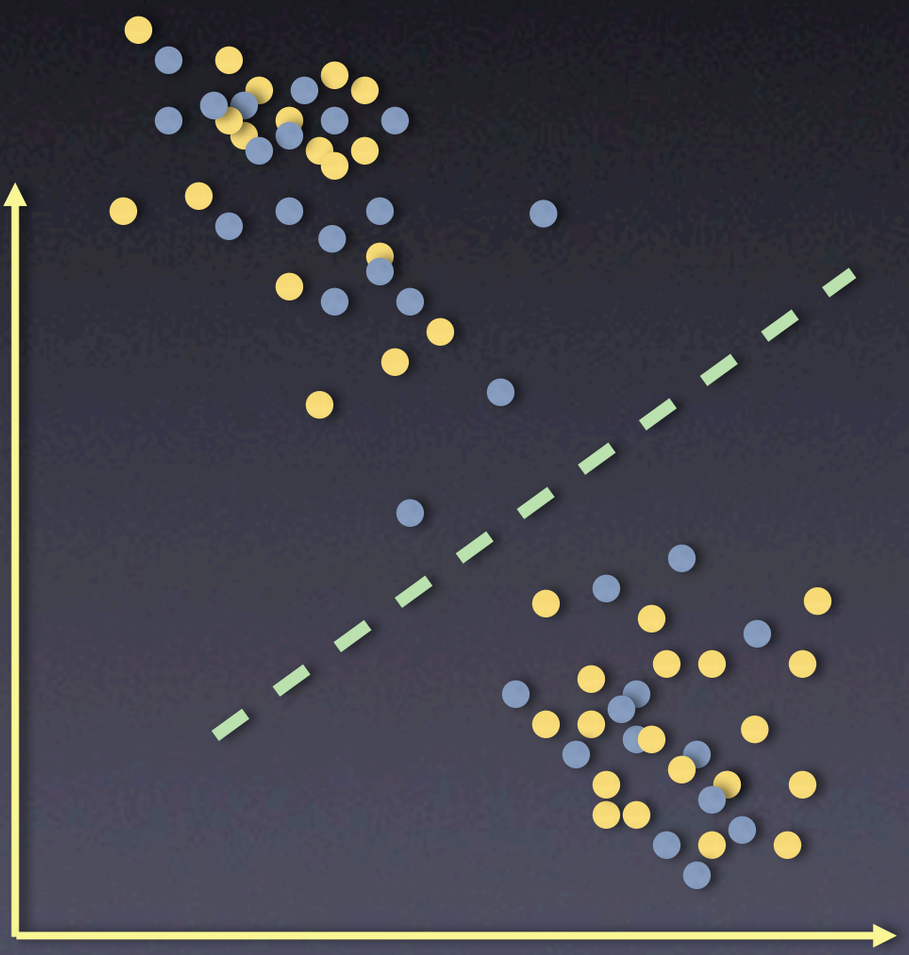
- Extracted phonological forms for the 5000 most frequent nouns and verbs
  - English: CELEX
  - French: BRULEX
- Identification cue: Word-initial onset and nucleus
- Category cue: Word-final nucleus and coda

# Discriminant Analysis (leave-one-out)



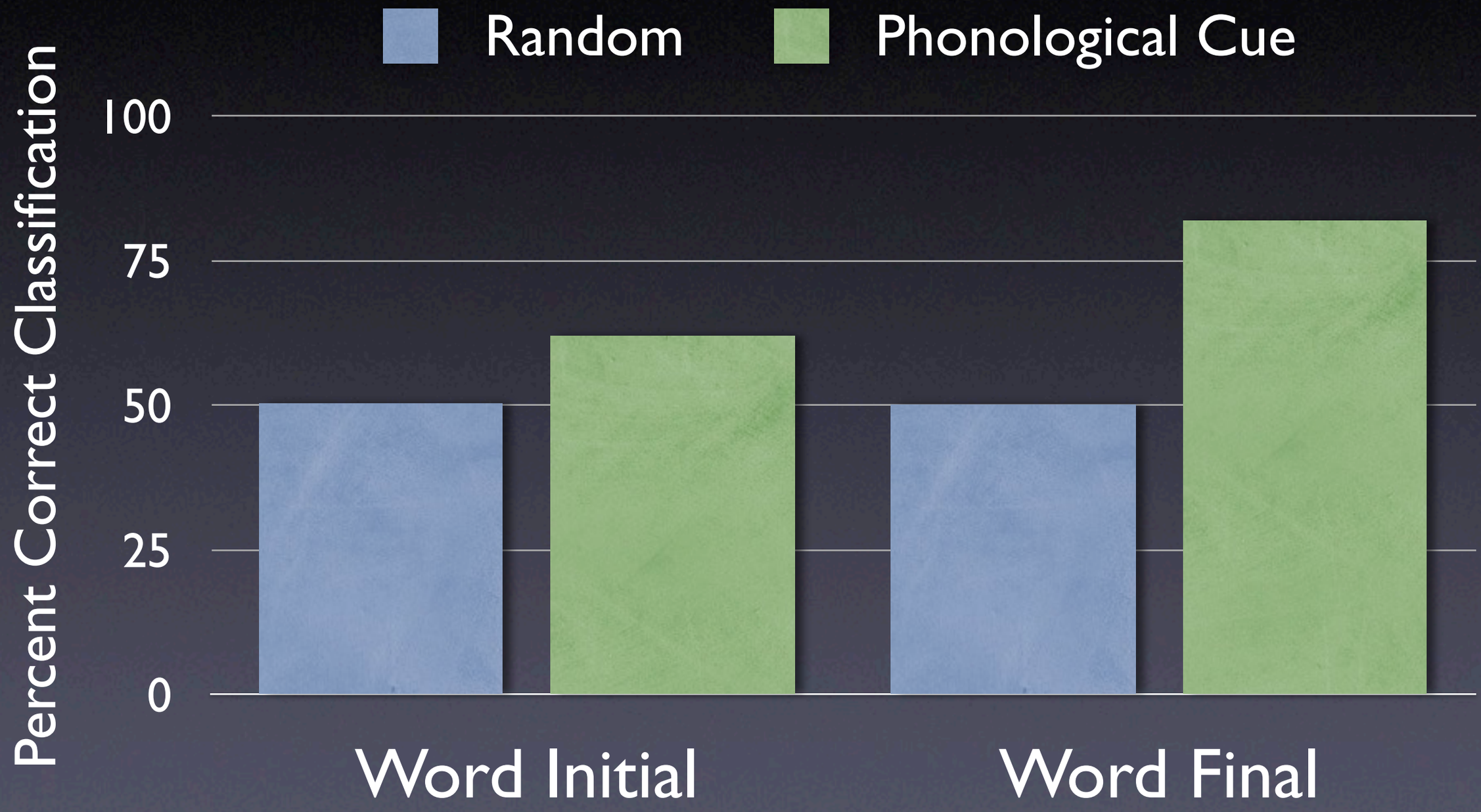

- Nouns
- Verbs

# Random Baseline




- Nouns
- Verbs

# Division of Labor in English Words




Source: Monaghan, Christiansen & Fitneva, in preparation

# Division of Labor in French Words

Source: Monaghan, Christiansen & Fitneva, in preparation

# Conclusion

- Language has evolved to balance arbitrariness and systematicity
  - Without context, arbitrariness impedes learning
  - With context, arbitrariness facilitates word individuation
- A division of labor between arbitrariness and systematicity within words maximizes learning of both individual meanings and lexical categories
- Vocabulary structure reflects multiple competing pressures on language acquisition

# Acknowledgments

Padraic Monaghan

Stanka Fitneva

# Thanks!