# Language Processing I

### The "Hub and Spokes" Framework Rogers & McClelland (2004)







## The "Hub and Spokes" Framework and the Brain





#### Area MT























## **Models of Semantics and Context (Control)**













#### "Categorize Item"





#### **"Judge Its Size"**









**Categorize Item:** Animal or Instrument?















Judge its size Bigger than a trash can?



Category Biased: Coherent covariation





Context Warping: Covariation based on size



#### Output "Sharpening"



## Semantic Attention Task

#### Task Verbally report size of object denoted by the word





Category

Mismatch

Category

## Semantic Attention Task

# TaskVerbally report size of object denoted by the word



### **Effect:**

#### Stable context (blocked) warps semantic space

Interleaved (no control)

"Size"

**Blocked (control)** "Animal Size" vs. "Instrument Size"









# Integrated Semantics and Control — Context Inference (ISC-CI)

Giallanza, Campbell, Rogers & Cohen (under review)





### Integrated Semantics and Control — Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)

### ISC model



semantic knowledge reflects the effects of statistical learning: coherent covariation



### Integrated Semantics and Control — Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)

### ISC model

#### A. ISC Model **Feature Label** is-black is-animal is-bird Context Dependent Context Context ndepend crow robin panther color **Context Labe** Item Label

#### semantic knowledge reflects the effects of statistical learning: *coherent covariation*

shaped in use by the influence of *context representations* that reflect explicit instructions, implicit behavior demands, and the other stimuli present or recently present in the environment



### Integrated Semantics and Control — Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)

### ISC model



- recently present in the environment

- semantic knowledge reflects the effects of statistical learning: coherent covariation

shaped in use by the influence of *context representations* that reflect explicit instructions, implicit behavior demands, and the other stimuli present or

context representations are themselves subject to the same mechanisms statistical learning, spanning multiple levels of abstraction, and driven by the statistics of behavioral affordances together with those of perception



### Integrated Semantics and Control – Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)

### ISC model

- semantic knowledge reflects the effects of statistical learning: coherent covariation



- shaped in use by the influence of *context representations* that reflect explicit instructions, implicit behavior demands, and the other stimuli present or recently present in the environment
- context representations are themselves subject to the same mechanisms statistical learning, spanning multiple levels of abstraction, and driven by the statistics of *behavioral affordances* together with those of perception

### • Still, many effects remain to be explained...



• Similarity Judgments:

- Order effects: asymetric similarity judgements

### • Similarity Judgments:

#### - Order effects: asymetric similarity judgements

How similar is a donkey to a horse?

### • Similarity Judgments:

#### - Order effects: asymetric similarity judgements

How similar is a donkey to a horse? Similar

### • Similarity Judgments:

#### - Order effects: asymetric similarity judgements

How similar is a donkey to a horse? Similar

How similar is a horse to a donkey?

### • Similarity Judgments:

#### - Order effects: asymetric similarity judgements

How similar is a donkey to a horse? Similar

How similar is a horse to a donkey? Less so

### • Similarity Judgments:

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so
- Multialternative effects: reversal of simiarlity judgments
## • Similarity Judgments:

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so

Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba

## • Similarity Judgments:

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so

Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba

## • Similarity Judgments:

Order effects: asymetric similarity judgements
How similar is a donkey to a horse? Similar
How similar is a horse to a donkey? Less so

### - Multialternative effects: reversal of simiarlity judgments

Which is Jamaica most similar to: England, Poland or Cuba Which is Jamaica most similar to: England, Russia or Cuba

## • Similarity Judgments:

Order effects: asymetric similarity judgements
How similar is a donkey to a horse? Similar
How similar is a horse to a donkey? Less so

### - Multialternative effects: reversal of simiarlity judgments

Which is Jamaica most similar to: England, Poland or Cuba Which is Jamaica most similar to: England, Russia or Cuba

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so
- Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba
  - Which is Jamaica most similar to: England, Russia or Cuba
- Triangle inequality effects: violations of transitive inference

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so
- Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba Which is Jamaica most similar to: England, Russia or Cuba
- Triangle inequality effects: violations of transitive inference [Nurse : Patient] :: [Mother : Child]

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so
- Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba Which is Jamaica most similar to: England, Russia or Cuba
- Triangle inequality effects: violations of transitive inference [Nurse : Patient] :: [Mother : Child] [Mother : Child] :: [Frog : Tadpole]

- Order effects: asymetric similarity judgements How similar is a donkey to a horse? Similar How similar is a horse to a donkey? Less so
- Multialternative effects: reversal of simiarlity judgments Which is Jamaica most similar to: England, Poland or Cuba Which is Jamaica most similar to: England, Russia or Cuba
- Triangle inequality effects: violations of transitive inference [Nurse : Patient] :: [Mother : Child]
  - [Mother : Child] :: [Frog : Tadpole]
  - [Nurse : Patient] :: [Frog : Tadpole]

- Similarity Judgments
- Category Judgements (Inductive Inference):

### • Similarity Judgments

• Category Judgements (Inductive Inference):

Given each of the following sets, which is more likely to be a member of the category:

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens or \{crows\} \rightarrow robins$ 

### Similarity Judgments

• Category Judgements (Inductive Inference):

Given each of the following sets, which is more likely to be a member of the category:

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$ 

### • Similarity Judgments

• Category Judgements (Inductive Inference):

Given each of the following sets, which is more likely to be a member of the category:

- Premise-conclusion similarity: *{crows} → ravens* or *{crows} → robins* 

- Conclusion typicality:

 $\{crows, ravens\} \rightarrow robins or \{crows, ravens\} \rightarrow penguins$ 

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity: *{crows} → ravens* or *{crows} → robins*
- Conclusion typicality: *{crows, ravens} → robins* or *{crows, ravens} → penguins*

### • Similarity Judgments

## • Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins \text{ or } \{crows, ravens\} \rightarrow penguins$
- Premise diversity:
  - $\{crows, robins\} \rightarrow sparrows or \{crows, ravens\} \rightarrow sparrows$

### • Similarity Judgments

## • Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins \text{ or } \{crows, ravens\} \rightarrow penguins$
- Premise diversity:
  - *{crows, robins} → sparrows* or *{crows, ravens} → sparrows*

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins \text{ or } \{crows, ravens\} \rightarrow penguins$
- Premise diversity:  $\{crows, robins\} \rightarrow sparrows or \{crows, ravens\} \rightarrow sparrows$
- In-category monotonicity  $\{crows, robins\} \rightarrow sparrows or \{crows\} \rightarrow sparrows$

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins \text{ or } \{crows, ravens\} \rightarrow penguins$
- Premise diversity:  $\{crows, robins\} \rightarrow sparrows or \{crows, ravens\} \rightarrow sparrows$
- In-category monotonicity (support spans query) *{crows, robins} → sparrows* or *{crows} → sparrows*

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity: {crows} → ravens or {crows} → robins
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins or \{crows, ravens\} \rightarrow penguins$
- Premise diversity: *{crows, robins} → sparrows* or *{crows, ravens} → sparrows*
- In-category monotonicity (support spans query)  $\{crows, robins\} \rightarrow sparrows or \{crows\} \rightarrow sparrows$
- In-category non-monotonicity *{brown bears, grizzly bears}* → *buffalo or {brown bears}* → *buffalo*

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins or \{crows, ravens\} \rightarrow penguins$
- Premise diversity: *{crows, robins} → sparrows* or *{crows, ravens} → sparrows*
- In-category monotonicity (support spans query)  $\{crows, robins\} \rightarrow sparrows or \{crows\} \rightarrow sparrows$
- **In-category non-monotonicity** (support is narrower than query) {brown bears, grizzly bears}  $\rightarrow$  buffalo or {brown bears}  $\rightarrow$  buffalo

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins$  or  $\{crows, ravens\} \rightarrow penguins$
- Premise diversity: *{crows, robins} → sparrows* or *{crows, ravens} → sparrows*
- In-category monotonicity (support spans query)  $\{crows, robins\} \rightarrow sparrows or \{crows\} \rightarrow sparrows$
- **In-category non-monotonicity** (support is narrower than query) *{brown bears, grizzly bears} → buffalo* or *{brown bears} → buffalo*
- Cross-category non-monotonicity  $\{flies, orangutans\} \rightarrow bees or \{flies\} \rightarrow bees$

### • Similarity Judgments

• Category Judgements (Inductive Inference):

- Premise-conclusion similarity:  $\{crows\} \rightarrow ravens \text{ or } \{crows\} \rightarrow robins$
- Conclusion typicality:  $\{crows, ravens\} \rightarrow robins$  or  $\{crows, ravens\} \rightarrow penguins$
- Premise diversity: *{crows, robins} → sparrows* or *{crows, ravens} → sparrows*
- In-category monotonicity (support spans query)  $\{crows, robins\} \rightarrow sparrows or \{crows\} \rightarrow sparrows$
- **In-category non-monotonicity** (support is narrower than query) *{brown bears, grizzly bears}*  $\rightarrow$  *buffalo* or *{brown bears}*  $\rightarrow$  *buffalo*
- **Cross-category non-monotonicity** (support is broader than query) *{flies, orangutans}*  $\rightarrow$  *bees* or *{flies}*  $\rightarrow$  *bees*

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories:
  - (Smith, Shobin, Rips, 1974; Miklov, 2013 - e.g., Word2Vec, ~ISC)

Metric Theories: *parametric distances* of representations in a high dimension vector space

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories:
  - (Smith, Shobin, Rips, 1974; Miklov, 2013 e.g. ,Word2Vec, ~ISC)
  - (Tversky, 1977; Osherson et al., 1990)

- Metric Theories: *parametric distances* of representations in a high dimension vector space

- Feature Contrast / Coverage Theories: *discrete (set theoretic) intersection* / overlap

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories:
  - (Smith, Shobin, Rips, 1974; Miklov, 2013 e.g. ,Word2Vec, ~ISC)
  - (Tversky, 1977; Osherson et al., 1990)
  - **Bayesian Inference Theories**: likelihood relative to prior (Xu & Tenenbaum, 2007; Griffiths et al., 2010)

- Metric Theories: *parametric distances* of representations in a high dimension vector space

- Feature Contrast / Coverage Theories: discrete (set theoretic) intersection / overlap

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories
- Context

- Warping of semantic space to change distances among items as a function of context

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories
- Context

  - **But**:
    - To what extent can context effects arise as emergent property of statistical learning of *continuous representations?*

- Warping of semantic space to change distances among items as a function of context

- Similarity Judgments
- Category Judgements (Inductive Inference)
- Theories
- Context

  - **But**:
    - To what extent can context effects arise as emergent property of statistical learning of *continuous representations*?
    - How do people *infer* this "on the fly?"

- Warping of semantic space to change distances among items as a function of context

### Integrated Semantics and Control – Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)







### Integrated Semantics and Control – Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)







### Integrated Semantics and Control — Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)

### Context





### Integrated Semantics and Control – Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)





# Integrated Semantics and Control — Context Inference (ISC-CI)

Giallanza, Campbell, Rogers & Cohen (under review)





### Integrated Semantics and Control — Context Inference (ISC-CI) Giallanza, Campbell, Rogers & Cohen (under review)



### • ISC-CI:



- relevant to that context: temporal autocorrelation
- and what other objects are likely to occur in that context

objects occurring together in a given context tend to share properties

these co-occurrence statistics are learned over the course of *development* 

this implicit knowledge provides a basis for inferring, from a few examples of objects encountered in a new context, both which features are relevant in that context


# Integrated Semantics and Control — Context Inference (ISC-CI)

Giallanza, Campbell, Rogers & Cohen (under review)







## Training

**Items from Leuven Concepts Database** (feature ratings) 







## • Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature





## Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)





## • Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)
- Predict which additional objects also likely to occur in that context {crow, robin, sparrow...}  $\rightarrow$  raven, ~jaguar





## Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)
- Predict which additional objects also likely to occur in that context {crow, robin, sparrow...}  $\rightarrow$  raven, ~jaguar

## Testing

- Tasks:

Similarity Judgements **Category Inference** 







## Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)

 $\rightarrow$ 

 $\rightarrow$ 

- Predict which additional objects also likely to occur in that context {crow, robin, sparrow...}  $\rightarrow$  raven, ~jaguar

## Testing

- Tasks: Similarity Judgements
  - Category Inference

### Format:

Argument (context): {support set} example members: {crow, robin}

<u>Conclusion</u> (stimulus) *{query set}* member of the category? sparrow?







## Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)
- Predict which additional objects also likely to occur in that context {crow, robin, sparrow...}  $\rightarrow$  raven, ~jaguar

## Testing

- Tasks:
  - Similarity Judgements Category Inference
- Format:

<u>Argument</u> (context): {support set} example members: {crow, robin}

**<u>Conclusion</u>** (stimulus) *{query set}* member of the category? sparrow?





## Training

- Items from Leuven Concepts Database (feature ratings)
- "Episodes" of 2 items sharing a single feature
- Predict context label for the one feature shared by all objects {crow, robin, sparrow...}  $\rightarrow$  Bird (context)
- Predict which additional objects also likely to occur in that context {crow, robin, sparrow...}  $\rightarrow$  raven, ~jaguar

## Testing

- Tasks:
  - Similarity Judgements Category Inference
- Format:

Argument (context): {support set} example members: {crow, robin}

Conclusion (stimulus) *{query set}* member of the category? sparrow?

**B. ISC-CI Model** Black **Attention /** Control mean Context







## Integrated Semantics and Control – Context Inference (ISC-CI)



# Integrated Semantics and Control – Context Inference (ISC-CI)



# Comparison of model and human choices in multi-alternative similarity judgments Tversky & Gati (1978)



Input

# Output

## Attention



### Embedding —

Input

Adda North

## Output



### Embedding

Add & Norm Multi-Head Attention ding Me Intending

## Output



nput 5,000 years worth of data!

### Embedding

## Output



nput 5,000 years worth of data!

# **Statistics** on Steroids!



Input

## Output







# Embedding

Output







# Embedding

Output

The

## Input









# Embedding









# Embedding









# Embedding

















**B. ISC-CI Model** 









**B. ISC-CI Model** 







### Same mechanisms of learning and representation of <u>statistical structure</u>

## Same mechanisms of learning and representation of <u>statistical structure</u>

 That structure ranges from simple (e.g., color vs. words) to **complex** (faces, scenes) and from *specific* (red, green... big, small...) to abstract (colors, size, living, self...)

- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from simple (e.g., color vs. words) to **complex** (faces, scenes) and from specific (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving action as perception

- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from simple (e.g., color vs. words) to complex (faces, scenes) and from specific (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving *action* as *perception*
- Control reflects its use to select information relevant in particular *contexts* and for particular *purposes*

- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from simple (e.g., color vs. words) to complex (faces, scenes) and from specific (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving action as perception
- Control reflects its use to select information relevant in particular contexts and for particular purposes
- Representations used for particular contexts and purposes may require *specialized processing apparatus* such as for rapid association formation, active maintenance, sequencing and search (*e.g., novel settings, and temporally extended goal directed action*)

- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from simple (e.g., color vs. words) to complex (faces, scenes) and from specific (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving action as perception
- Control reflects its use to select information relevant in particular contexts and for particular purposes
- Representations used for particular contexts and purposes may require *specialized processing apparatus* such as for rapid association formation, active maintenance, sequencing and search (*e.g., novel settings, and temporally extended goal directed action*)
- (Attentional) control: a parametric warping of semantic space by context
- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from simple (e.g., color vs. words) to complex (faces, scenes) and from specific (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving action as perception
- Control reflects its use to select information relevant in particular contexts and for particular purposes
- Representations used for particular contexts and purposes may require *specialized processing apparatus* such as for rapid association formation, active maintenance, sequencing and search (*e.g., novel settings, and temporally extended goal directed action*)
- (Attentional) control: a parametric warping of semantic space by context
- Semantics and control are inextricably intertwined

- Same mechanisms of learning and representation of <u>statistical structure</u>
- That structure ranges from **simple** (e.g., color vs. words) to **complex** (faces, scenes) and from *specific* (red, green... big, small...) to abstract (colors, size, living, self...)
- It just as sensitive to statistical relationships involving action as perception
- Control reflects its use to *select information* relevant in particular *contexts* and for particular *purposes*
- Representations used for particular contexts and purposes may require specialized processing apparatus such as for rapid association formation, active maintenance, sequencing and search (e.g., novel settings, and temporally extended goal directed action)
- Semantics and control are inextricably intertwined

# • (Attentional) control: a parametric warping of semantic space by context



Connectionist networks are an appropriate "tool box" for studying such effects:

#### • statistical learning mechanisms that

- capture relevant semantic structure through experience
- that reflects relationships present in the *environment*
- under pressure of their usefulness for prediction and/or action

- <u>Connectionist networks</u> are an appropriate "tool box" for studying such effects:

- statistical learning mechanisms that
  - capture relevant semantic structure through experience
  - that reflects relationships present in the environment
  - under pressure of their usefulness for *prediction* and/or *action*
- mechanisms for context processing and control that
  - help exploit and shape representational structure (e.g., via biasing effects and non-linearities)
  - accommodate special processing requirements for different purposes (e.g., episodic memory mechanisms for rapid association formation and gated attractors [LSTMs] for active maintenance, sequencing and search

<u>Connectionist networks</u> are an appropriate "tool box" for studying such effects: