Scalar implicature:

a whirlwind tour with stops in processing, development and disorder

Jesse Snedeker
Harvard University
Tubingen
July 2015

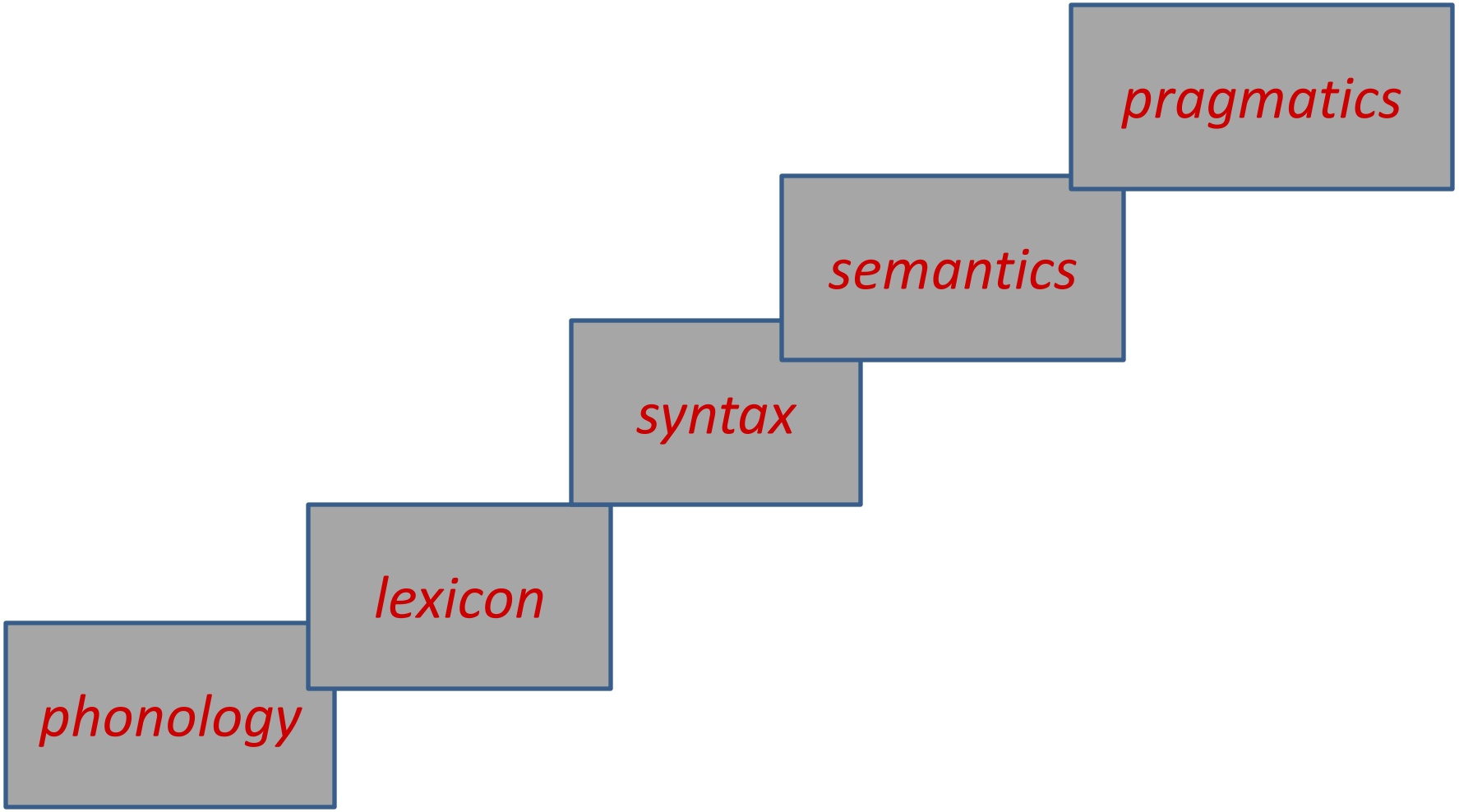
Yi Ting Huang

Outline

- Grounding assumptions
- A modest proposal
- The data behind the proposal
  1. Implicature typically takes time and effort
  2. Instant SI's occur only when pre-encoding is plausible
  3. SI proficiency develops slowly
  4. In disordered populations SI patterns with language ability

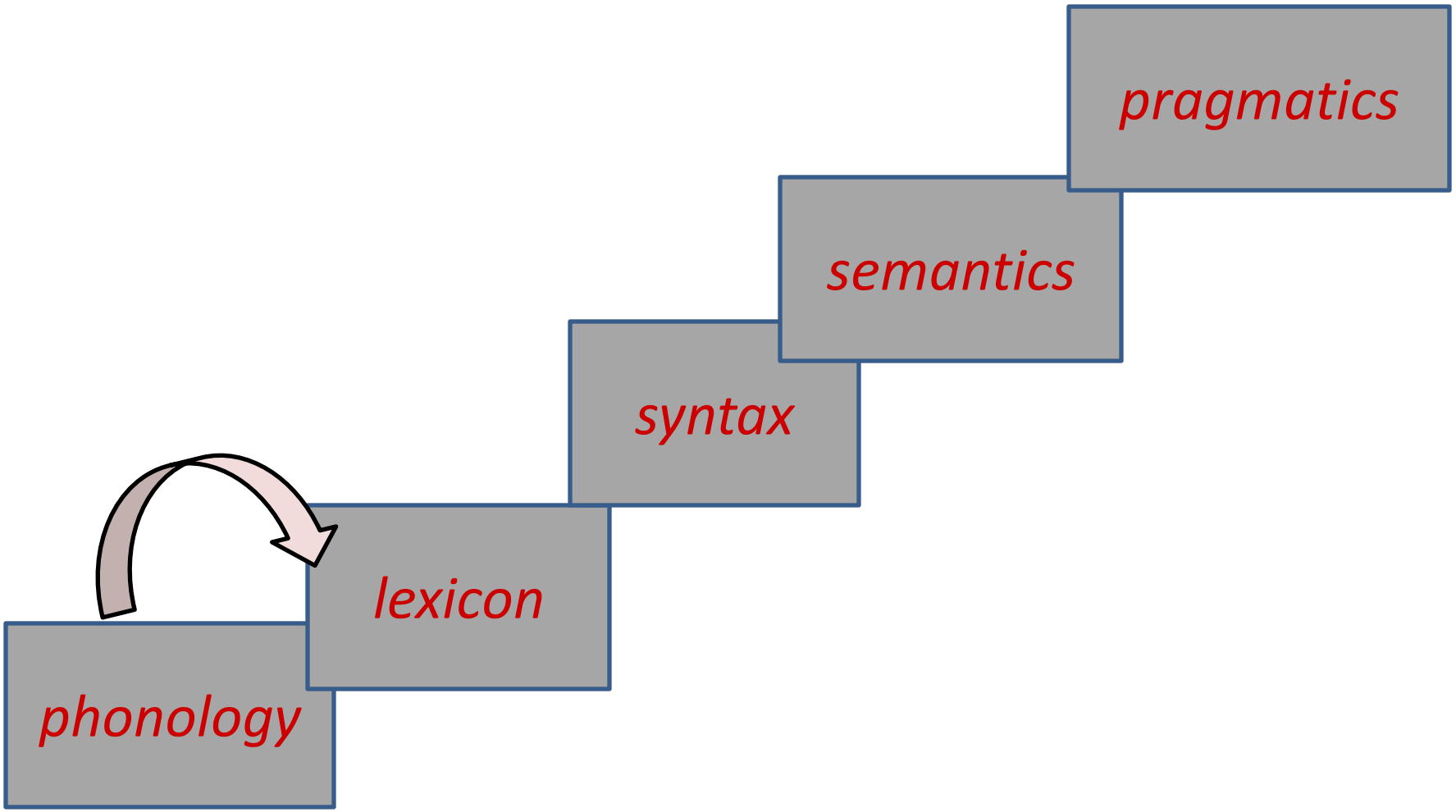
21st century standard model

1. Comprehension builds a partially ordered series of representations



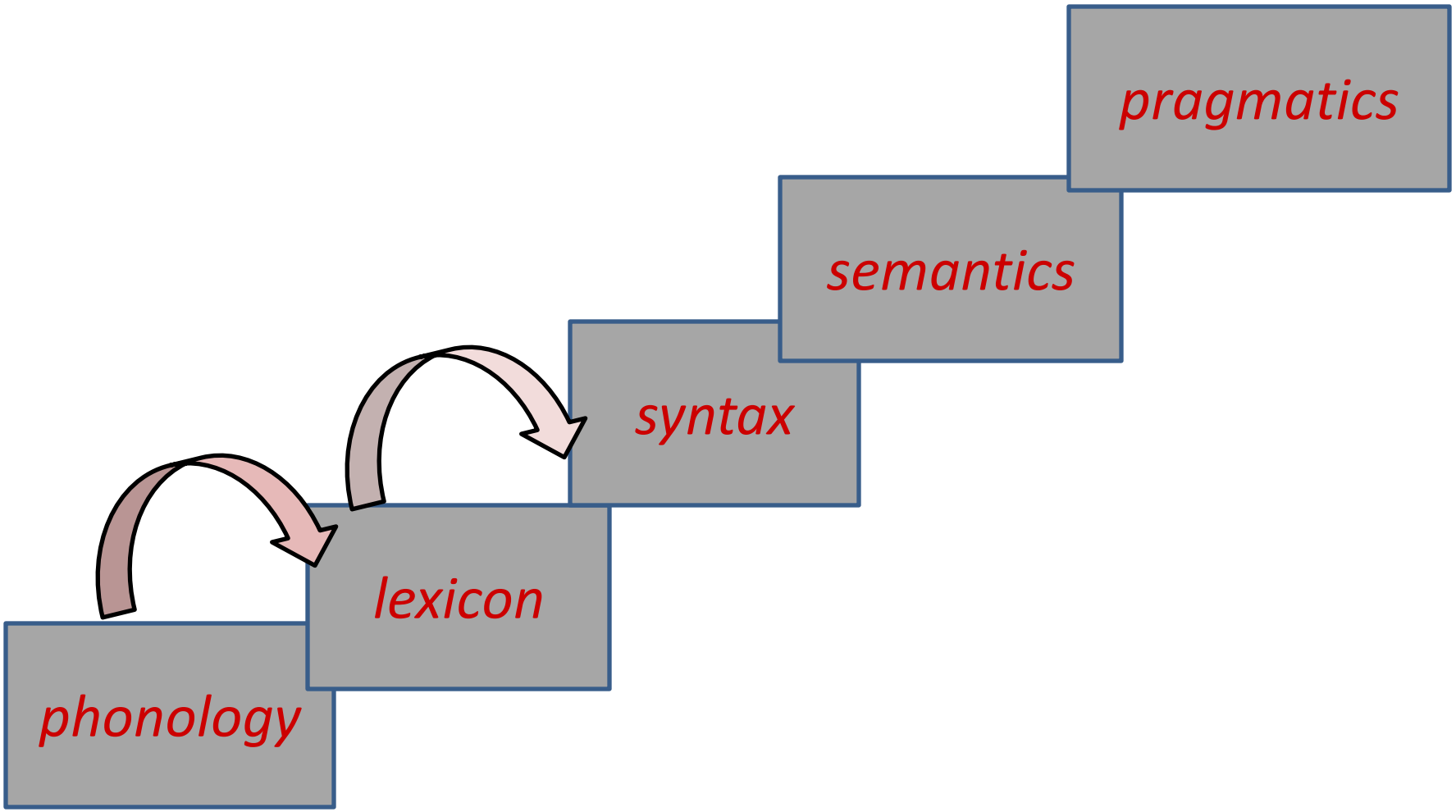
21st century standard model

2. Cascaded processing (incrementality)



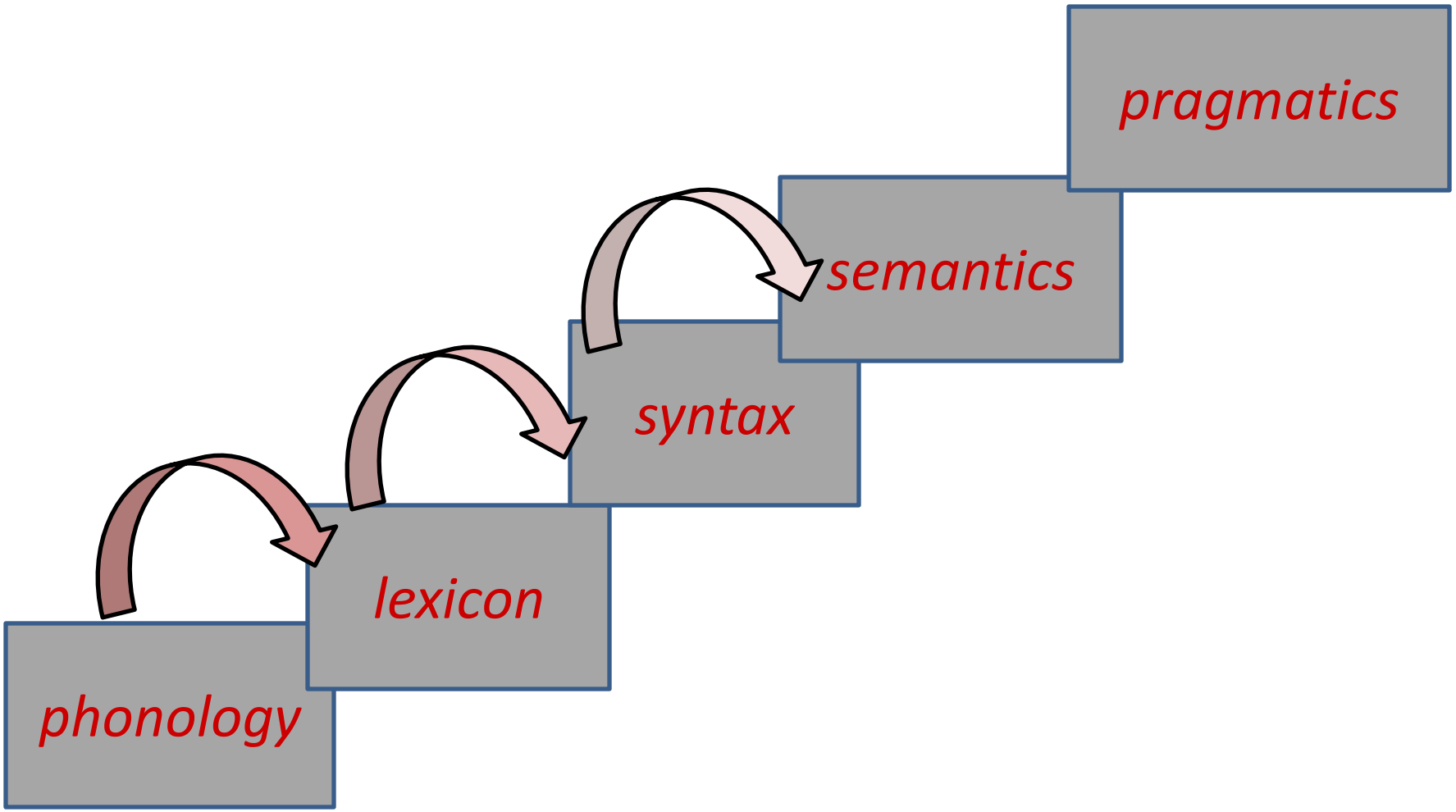
21st century standard model

2. Cascaded processing (incrementality)



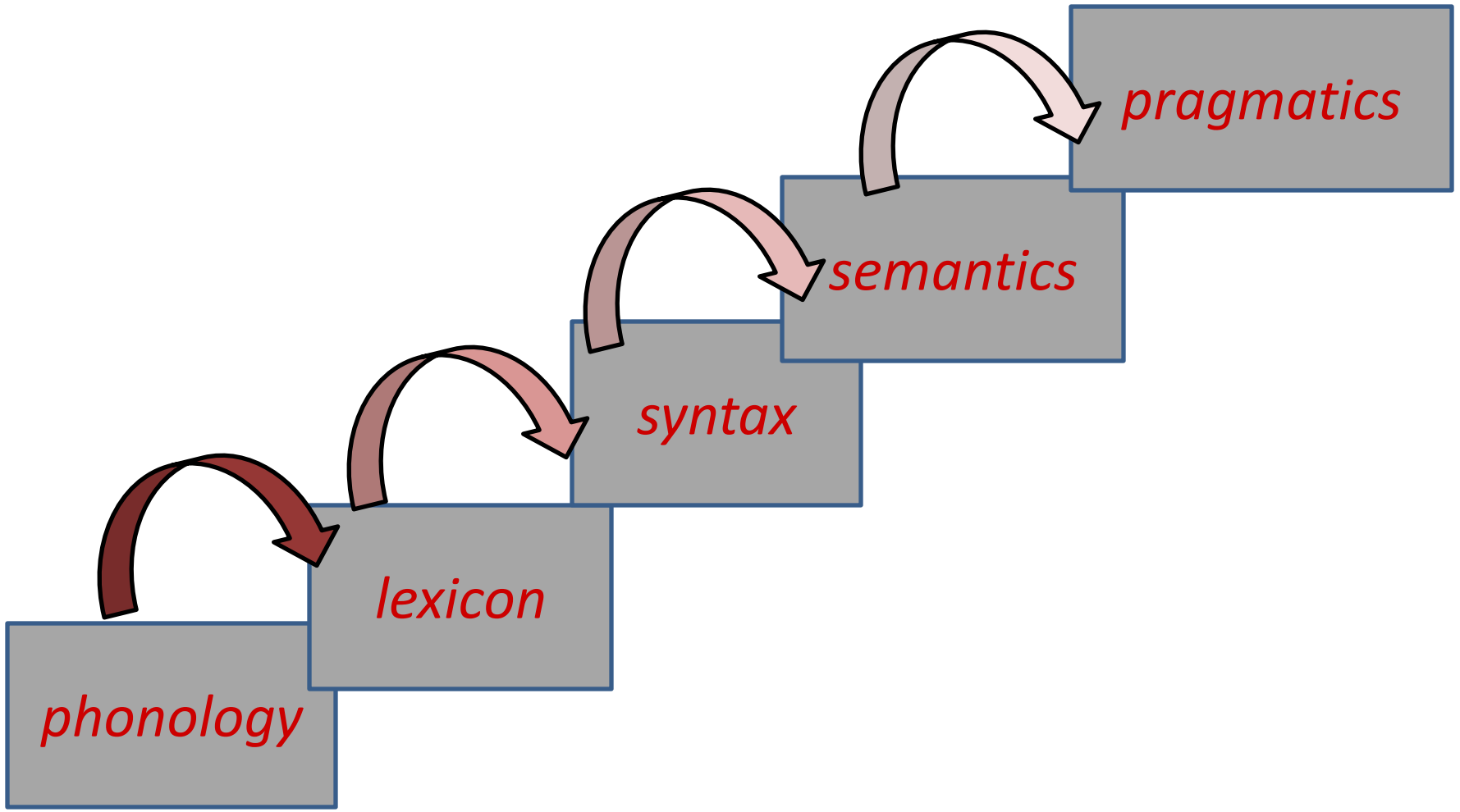
21st century standard model

2. Cascaded processing (incrementality)



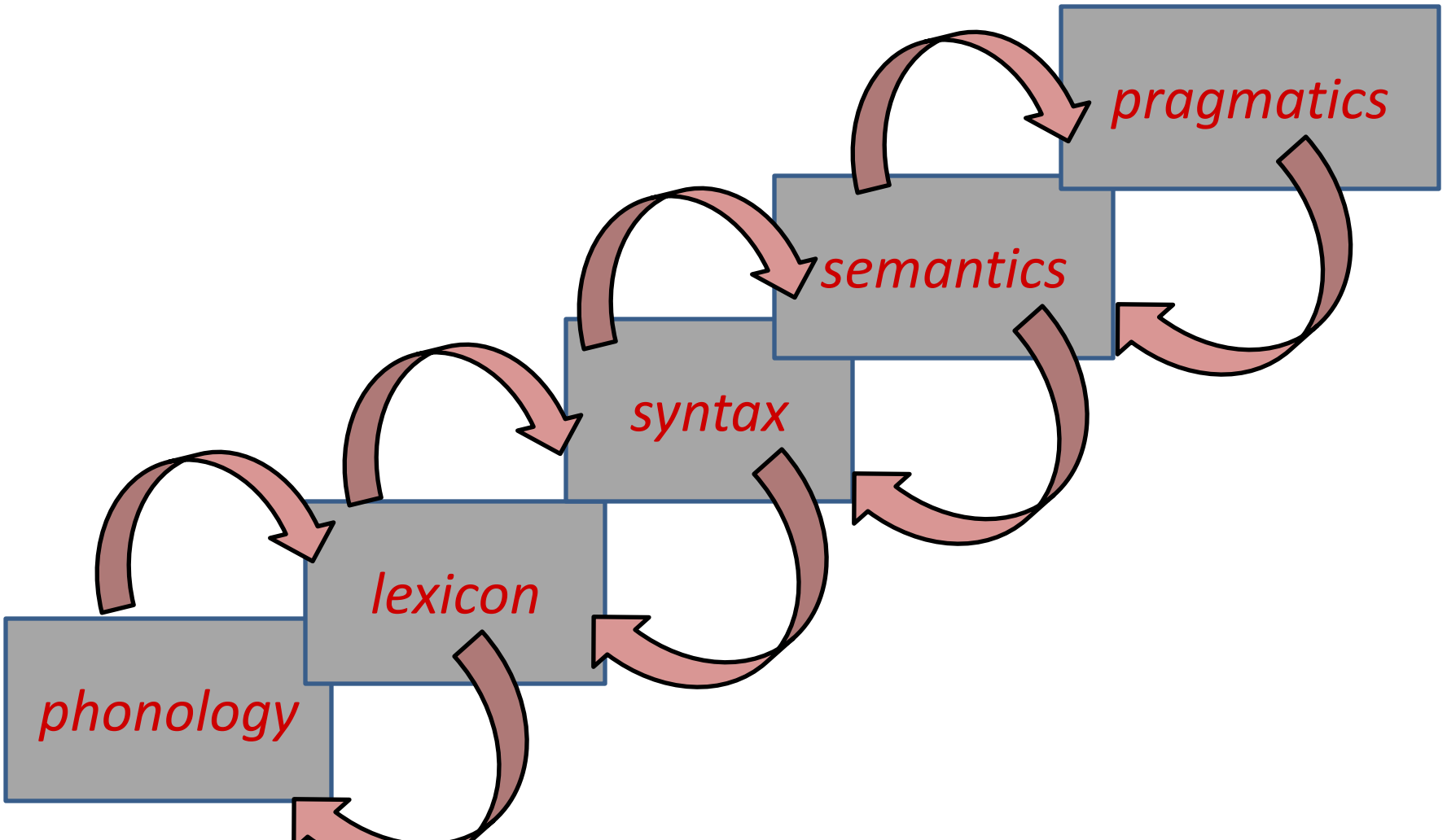
21st century standard model

2. Cascaded processing (incrementality)



21st century standard model

3. Processing is interactive: both directions

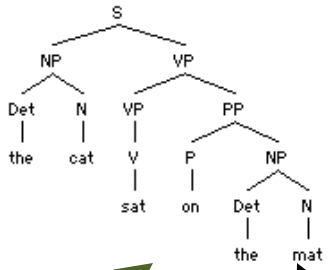


3. Interactivity: many inputs

pragmatic interpretation

semantics

$\exists$ x [cat (x) $\wedge$ on mat (x)]
$\wedge \forall$ y [cat (y) $\wedge$ on mat (y)] x=y



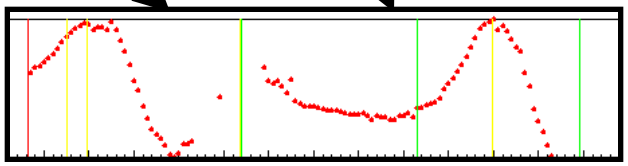
syntax

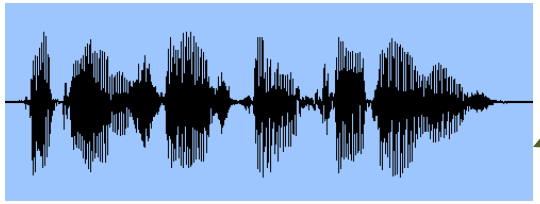
lexicon

/kæt/: noun, singular, animate
/sæt/: verb, past, intransitive
/mæt/: noun, singular, inanimate

phonology

əkætsætmæt



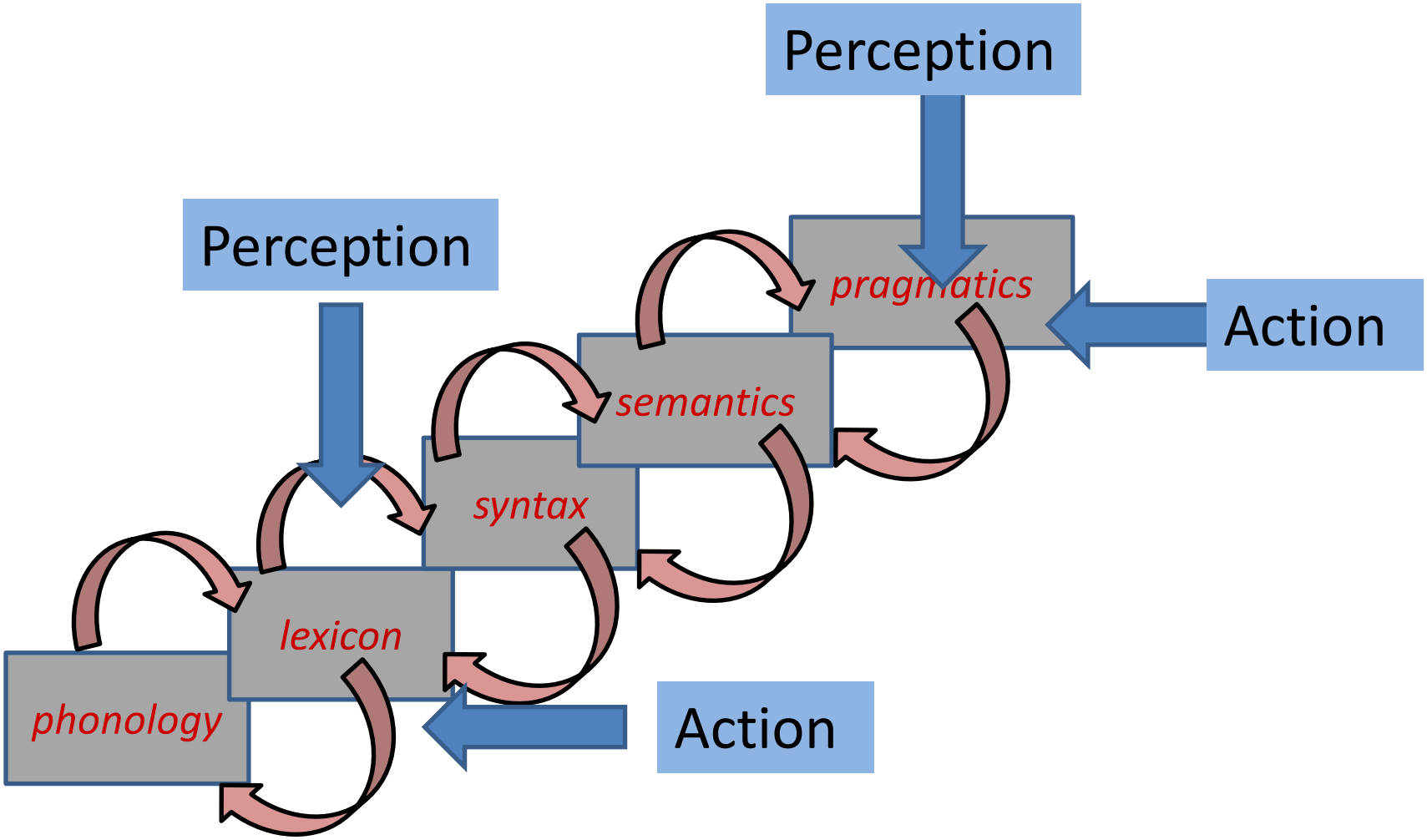


acoustic processing

prosody

21st century standard model

4. No walls around language



21st Century Standard Model

1. Levels of representation
2. Incremental
3. Interactive
   - Corollary: under many circumstances processing will be predictive
4. In contact with perception and action
   - Corollary: introduces the possibility of top-down prediction of speech

- Does this mean that all natural inferences are made instantly, with no delay?
- Of course not, cognitive operations unfold over time
  - Can be done ahead of time
  - Can be stored
  - But they are not atemporal

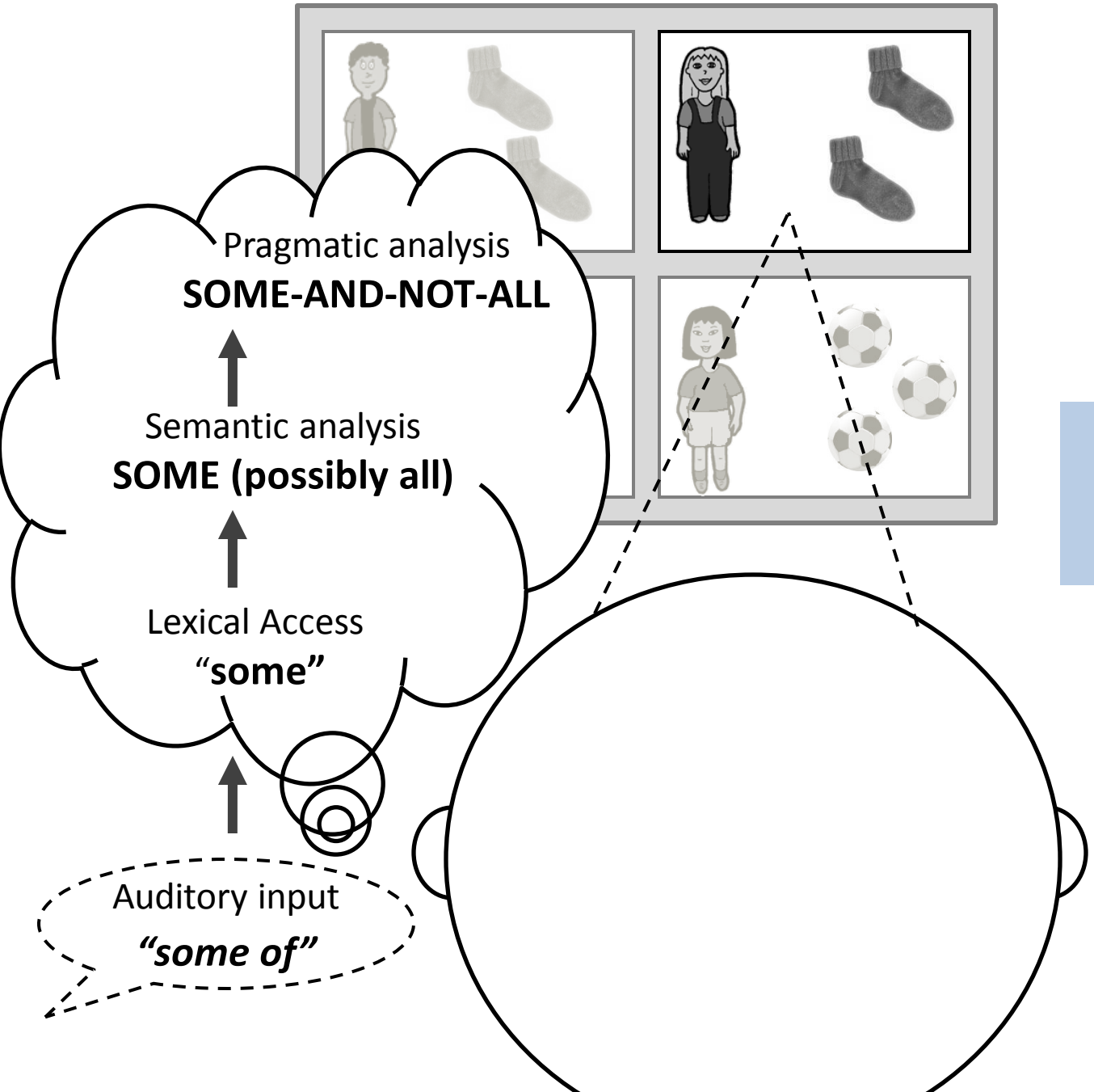
How are scalar implicatures calculated?

Bottom-up

- Hear “some”
- Retrieve its meaning
- Activate stronger alternative (*all*)
- Construct enriched meaning
- Evaluate / link to context

Dependent on context!

Remember, this is incremental and interactive
(we reject the “2-stage” label)

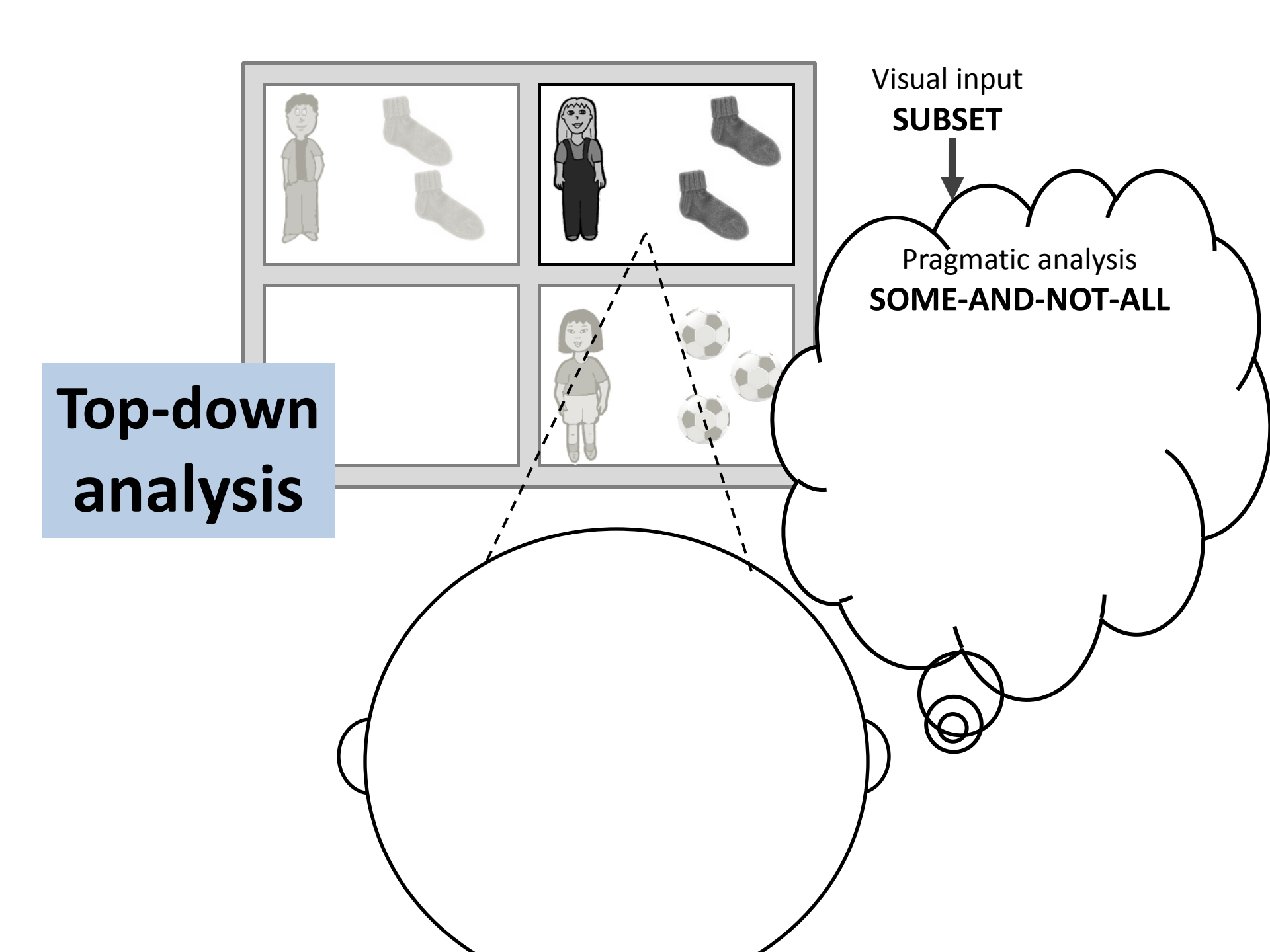


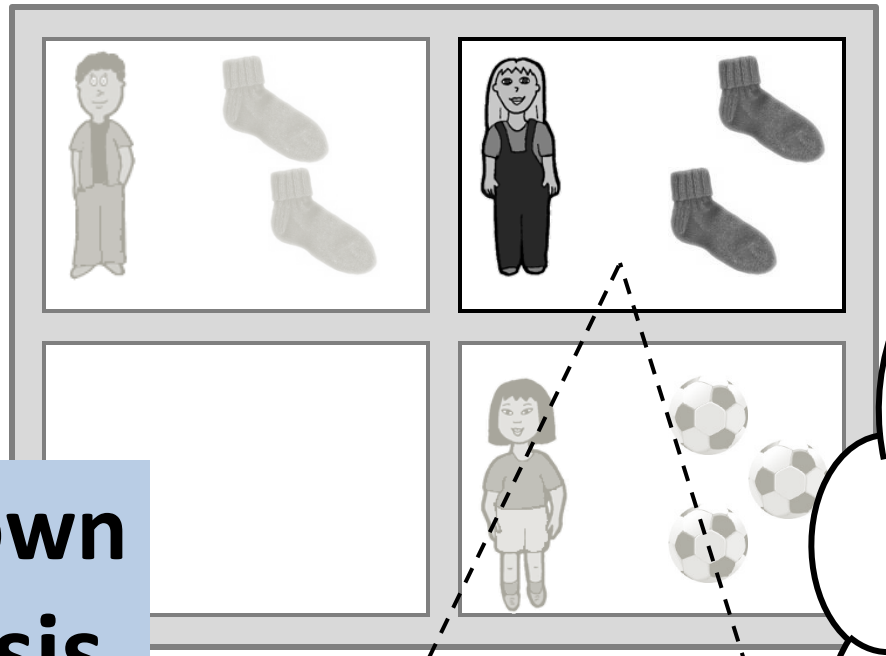
Bottom-up analysis

How are scalar implicatures calculated?

Top-down

- Listener sees display (knows the situation)
- Encodes a “message level” representation of possible referents (GIRL + SUBSET OF X’S)
- Begins to link to lower levels of representation (semantic, maybe even lexical)


Visual input
SUBSET
Pragmatic analysis
SOME-AND-NOT-ALL
Top-down analysis



Visual input

SUBSET

Pragmatic analysis

SOME-AND-NOT-ALL

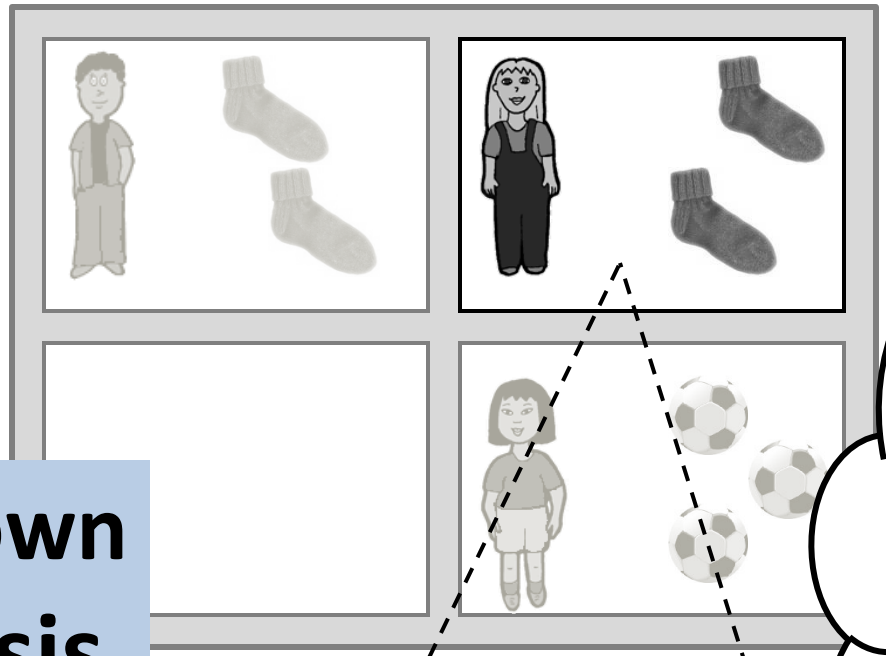
Semantic Analysis

SOME

Lexical access

“**some**”

Top-down analysis



Visual input
SUBSET

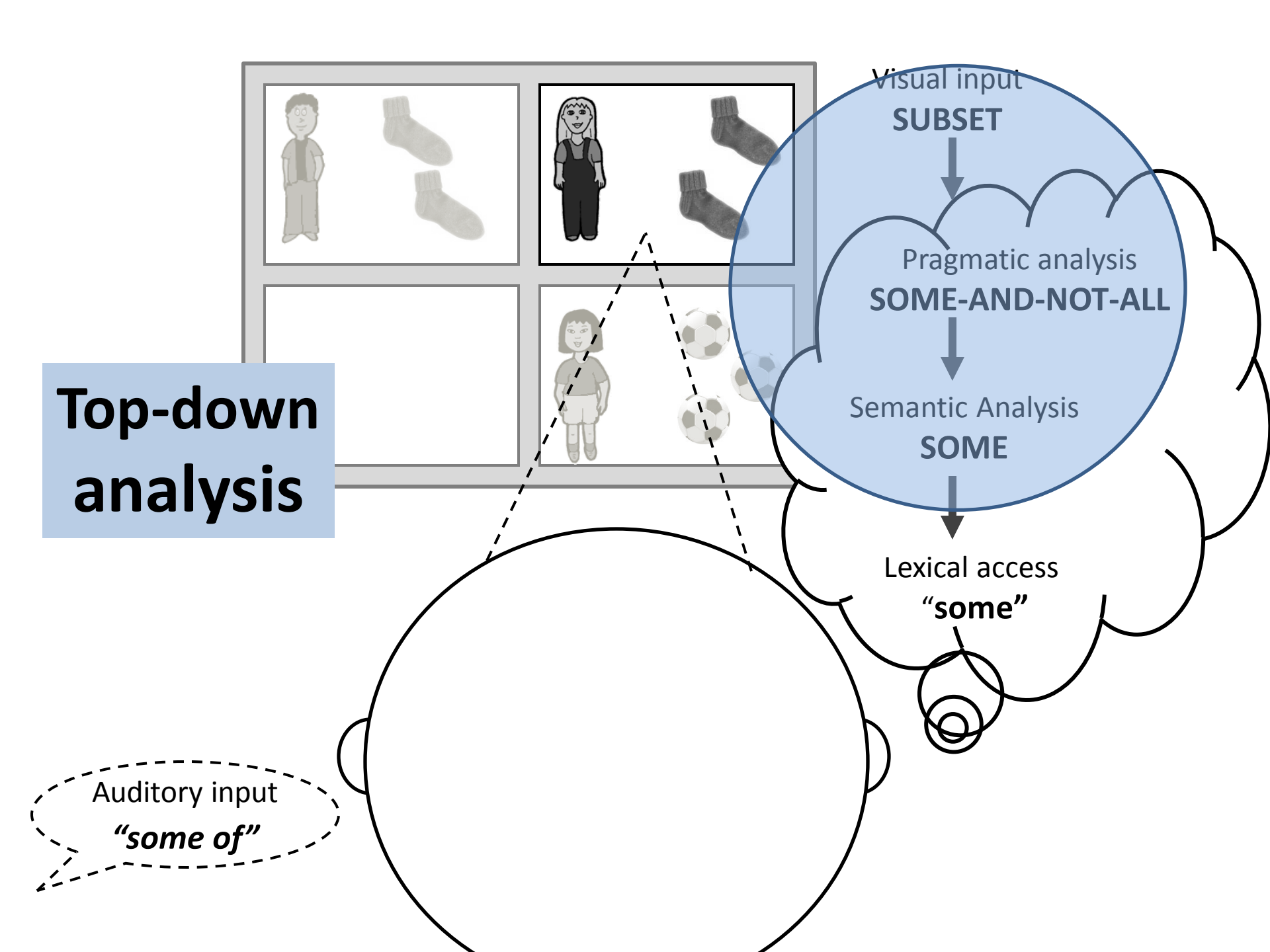
Pragmatic analysis
SOME-AND-NOT-ALL

Semantic Analysis
SOME

Lexical access
“**some**”

Top-down analysis

Auditory input
“some of”


Top-down analysis
Visual input
SUBSET
Pragmatic analysis
SOME-AND-NOT-ALL
Semantic Analysis
SOME
Lexical access
"some"
Auditory input
"some of"

Predictions

- Bottom-up
  - Scalar upper bound delayed relative to lexically encoded upper and lower bounds
  - Occurs when verbal encoding is difficult
    - Messages more unpredictable to comprehender
    - Multiple construals of given referent
- Top-down
  - Scalar upper bound guide reference resolution as rapidly as lexical bounds
  - Occurs when a verbal encoding is easy
    - Facts already known to listener (visual world)
    - Single salient construal of each referent in task

Outline

- Grounding assumptions
- A modest proposal
- The data behind the proposal
  1. Implicature typically takes time and effort
  2. Instant SI's occur only when pre-encoding is plausible
  3. SI proficiency develops slowly
  4. In disordered populations SI patterns with language ability

SI's typically require time and effort

1. Sentence judgment studies

   Bott & Noveck, 2004; Bott, Bailey & Grodner, 2012; Marty & Chemla, 2011

2. Dual-task studies

   DeNeys & Schaeken, 2007; Dieussaert, Verkerk, Gillard & Schaeken, 2011; Marty & Chemla, 2011; Marty & Chemla, 2011

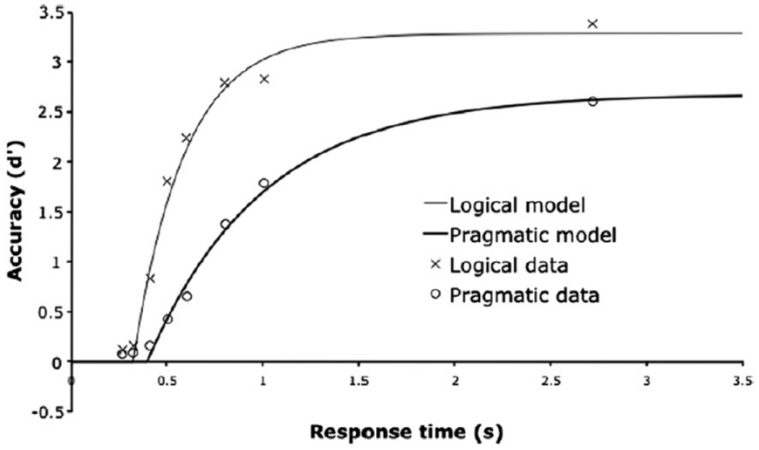
3. Reading studies

   Breheny, Katsos & Williams, 2006; Bergen & Grodner, 2010; Hartshorne & Snedeker, still under review; Nieuwland, Dittman & Kuperberg , 2010

Judgment tasks:
Bott, Bailey & Grodner (2012)

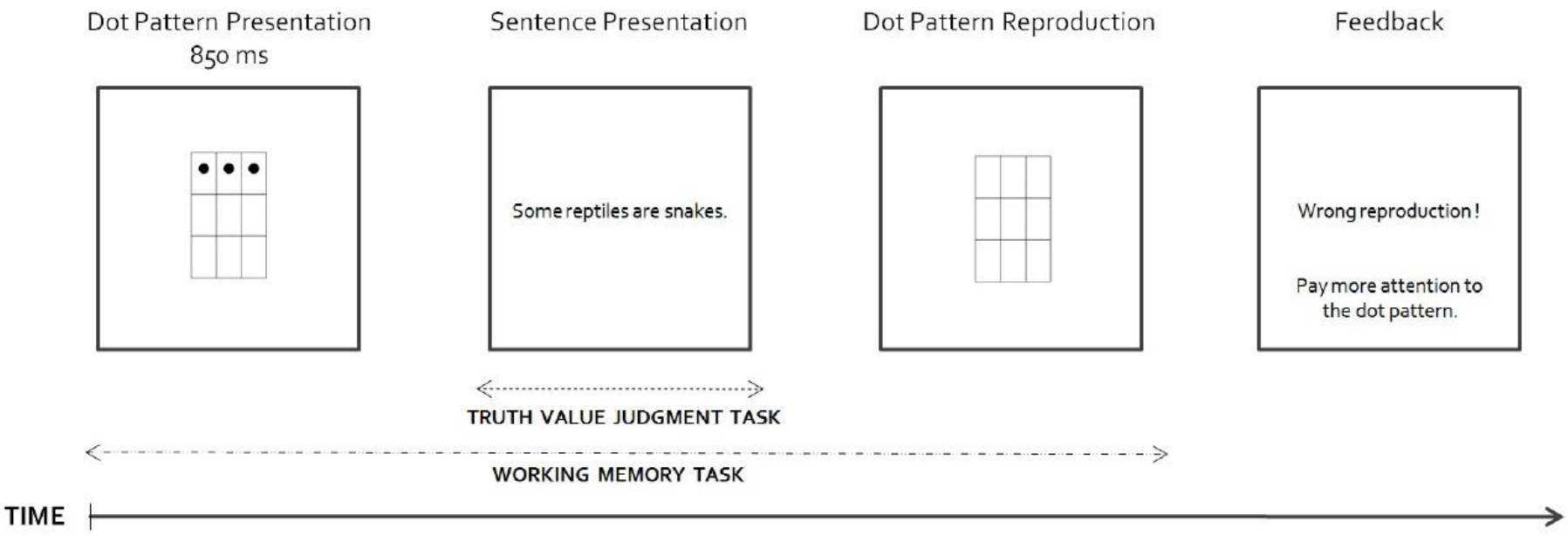
- Speeded verification of underinformative sentences (SAT task)
- Delay for calculating SI
- Not due to speed accuracy tradeoff: shift in starting point and slope
- Not merely verification: pragmatic “some” slower than “only some”

“Some elephants are mammals”



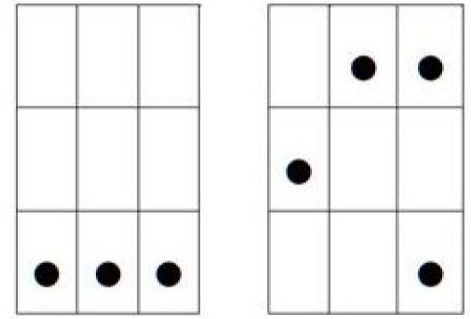
Dual-task paradigm

- Cognitive load reduces calculation of scalar implicatures (DeNeys & Schaeken, 2007; Dieussaert, Verkerk, Gillard & Schaeken, 2011; Marty & Chemla, 2011; Marty, Chemla & Spector, 2011)

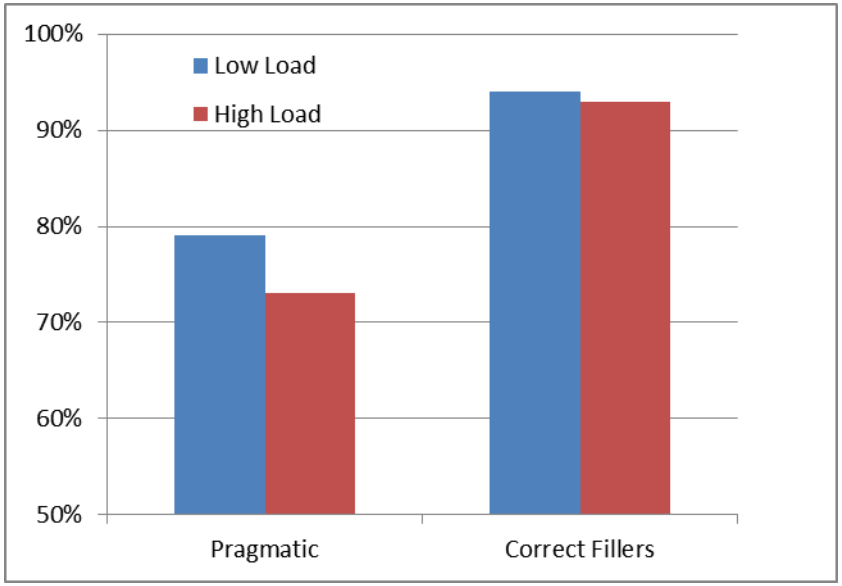


Dual-task paradigm

- Cognitive load reduces calculation of scalar implicatures (DeNeys & Schaeken, 2007; Dieussaert, Verkerk, Gillard & Schaeken, 2011; Marty & Chemla, 2011; Marty, Chemla & Spector, 2011)



Low load High load



Data from: DeNeys & Schaeken (2007)

Dual-task paradigm

- Cognitive load reduces calculation of scalar implicatures (DeNeys & Schaeken, 2007; Dieussaert, Verkerk, Gillard & Schaeken, 2011; Marty & Chemla, 2011; Marty & Chemla, 2011)
- Load does not reliably interfere with semantic upper bounds ("only some") (Marty & Chemla, 2011)
- Opposite effect for numbers (Marty, Chemla & Spector, 2011)

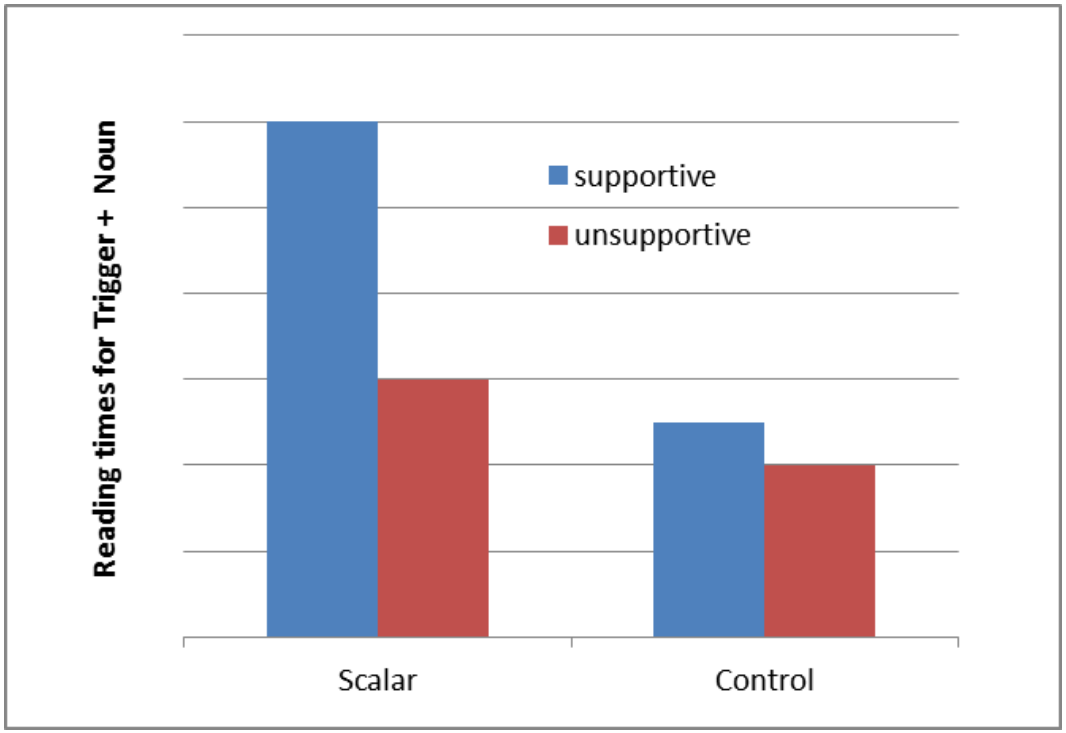
Reading time experiments

Recipe (from Breheny et al., 2006, illustrated with Bergen & Groder, 2012)

- Contexts
  - Supportive: “Before the hurricane landed, I checked every house in town.”
  - Non-supportive: “Before the hurricane landed, I volunteered to help out in town.”
- Trigger
  - Scalar: “Some of the residents had evacuated”
  - Control: “Only some of the residents had evacuated”
- Anaphor (probes upper bound)
  - “The rest stayed home and foolishly risked their lives”

Reading time experiments

- Slow down at trigger for Scalar in supportive contexts*



Data from Bergen & Grodner

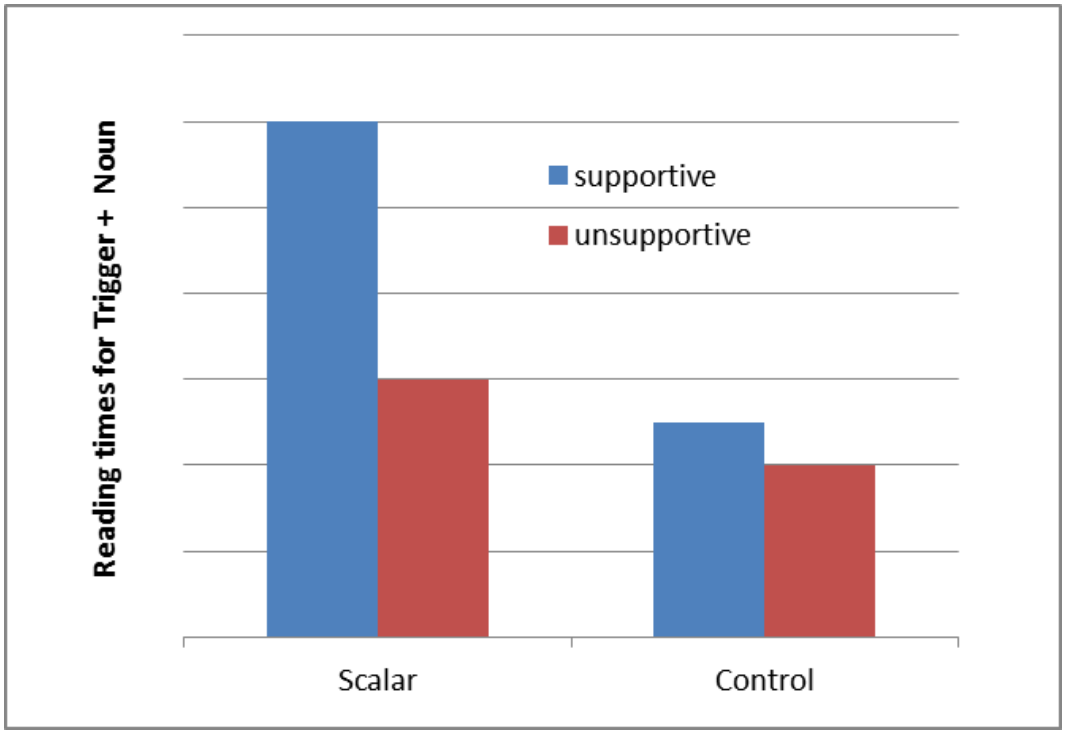
* But see Hartshorne & Snedeker for caveats

Reading time experiments

- Slow down at trigger for Scalar in supportive contexts*

Interpretation:

- SI takes effort
- Effort begins immediately
- But only when context calls it up

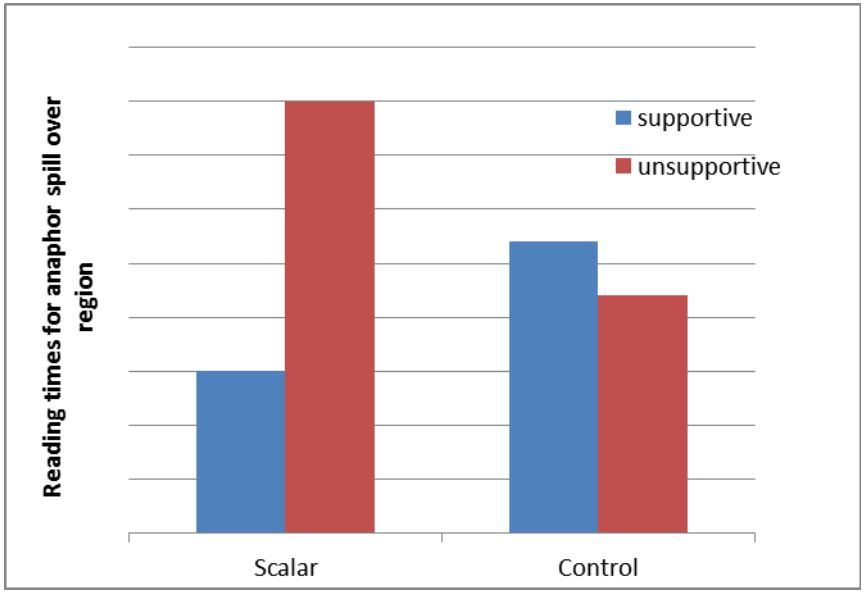


Data from Bergen & Grodner

* But see Hartshorne & Snedeker for caveats

Reading time experiments

- Slow down after anaphor for scalars in unsupportive contexts



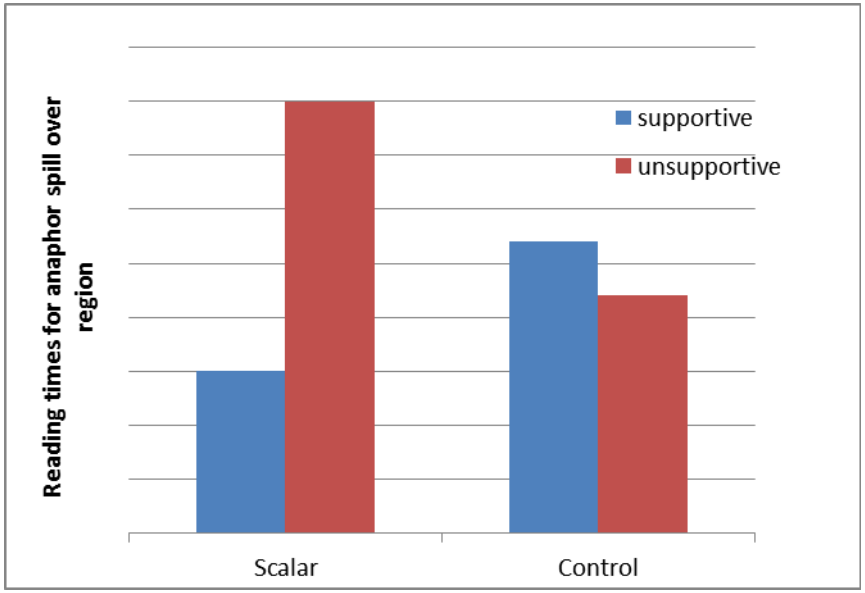
Data from Bergen & Grodner

Reading time experiments

- Slow down after anaphor for scalars in unsupportive contexts

Interpretation:

- Upper bound calculated in supportive contexts and controls
- Not in unsupportive contexts



Data from Bergen & Grodner

Reading time experiments

- How fast is that upper bound calculated?
- Mean time from trigger to anaphor effect
  - Bergen & Grodner: ~2,400 ms
  - Breheny et al: ~2000 ms
  - Nieuwland et al: ~1700 ms

Reading time experiments

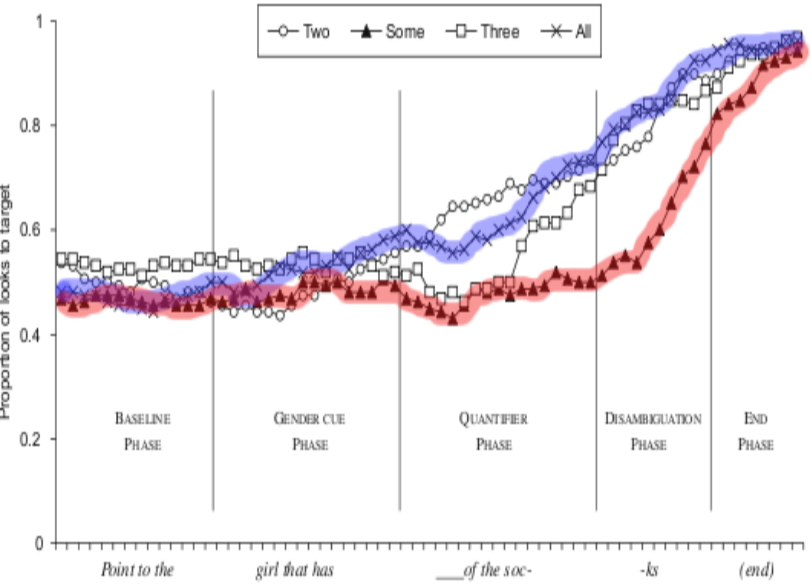
- How fast is that upper bound calculated?
- Mean time from trigger to anaphor effect
  - Bergen & Grodner: ~2,400 ms
  - Breheny et al: ~2000 ms
  - Nieuwland et al: ~1700 ms
- Hartshorne & Snedeker manipulate distance
  - No anaphor effect at 1500 ms
  - Robust anaphor effect at 3000 ms
  - Adding upper bound takes time

Outline

- Grounding assumptions
- A modest proposal
- The data behind the proposal
  1. Implicature typically takes time and effort
  2. Instant SI's occur only when pre-encoding is plausible
  3. SI proficiency develops slowly
  4. In disordered populations SI patterns with language ability

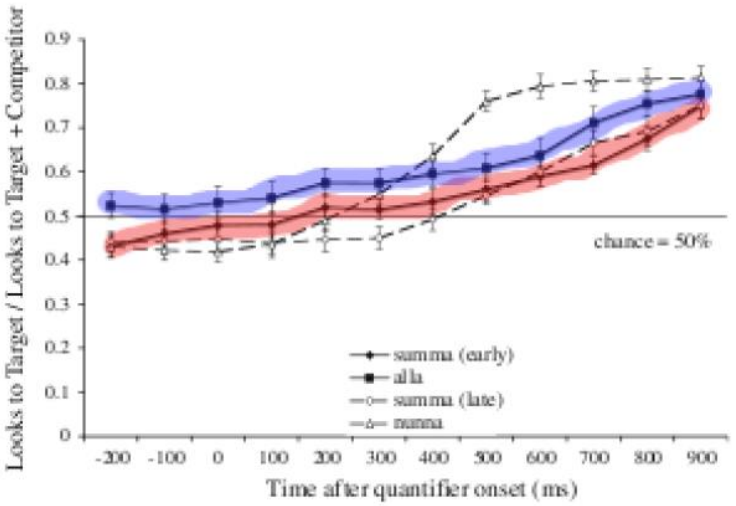
Divergent Findings in Visual World Paradigm

Delayed Upper Bound for “Some”



Huang & Snedeker (2009)

Instant Upper Bound for “Some”

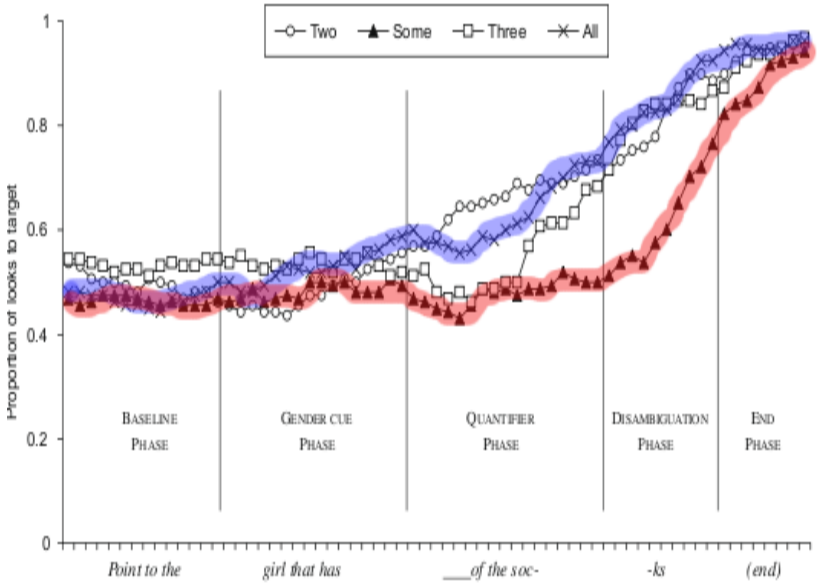


Grodner et al. (2010)

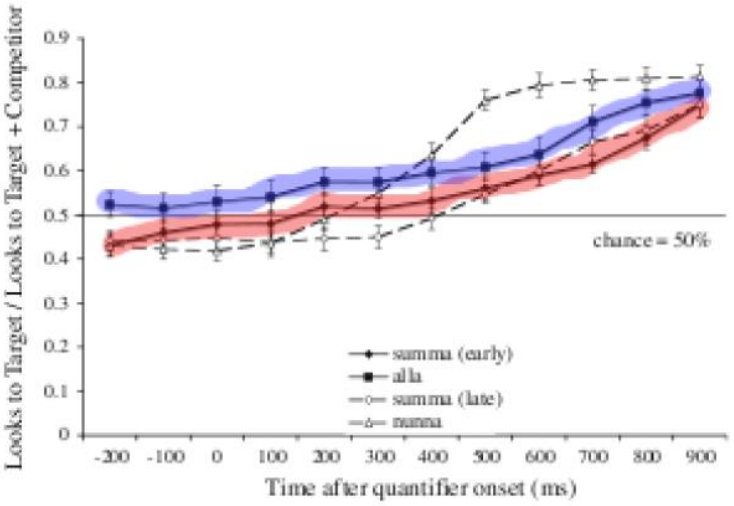
Methodological differences

- Pronunciation “summa” vs. some of
- Embedded in stories vs. not
- Length of experiment
- Number trials (Huang, Hahn & Snedeker; Degen & Tanenhaus)

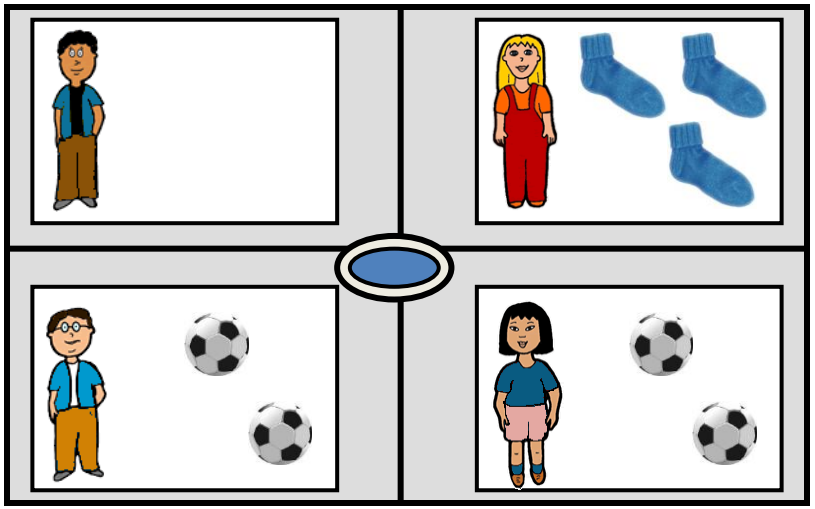
Huang & Snedeker (2009)



Grodner et al. (2010)



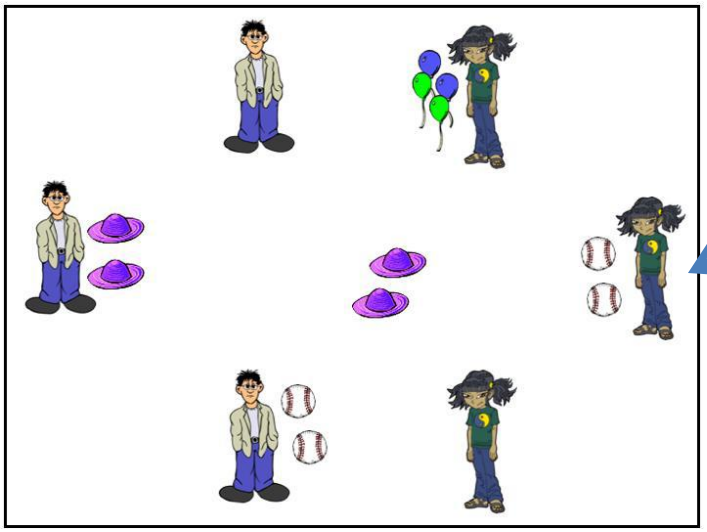
Comparison of studies



Dual Encoding:

The girl with some of the soccer balls
The girl with two of the soccer balls.

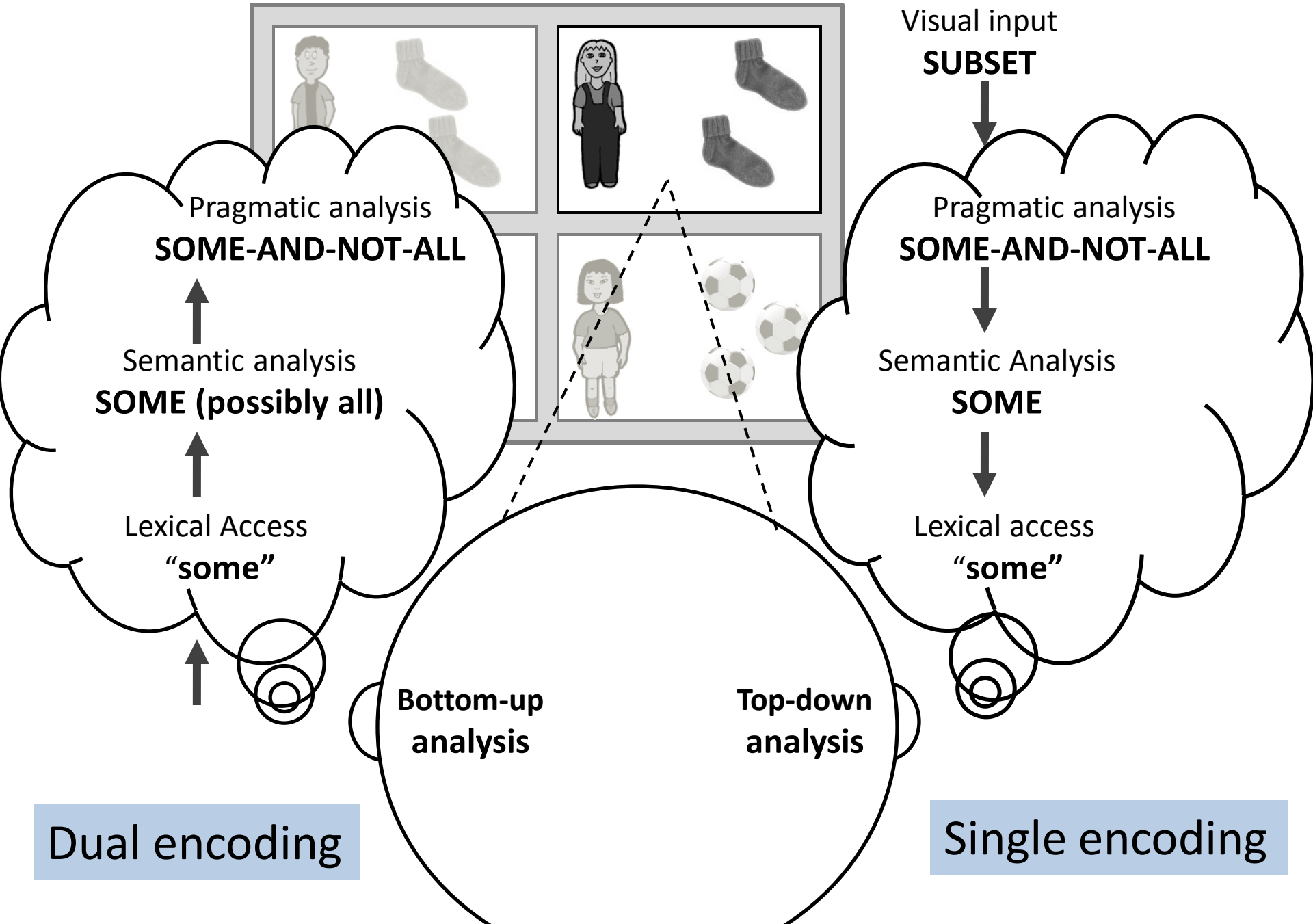
SI delayed



Single Encoding:

The girl with some of the balls

Immediate SI


Visual input
SUBSET
Pragmatic analysis
SOME-AND-NOT-ALL
Semantic analysis
SOME (possibly all)
Lexical Access
"some"
Pragmatic analysis
SOME-AND-NOT-ALL
Semantic Analysis
SOME
Lexical access
"some"
Bottom-up analysis
Top-down analysis
Dual encoding
Single encoding

Robust generalization across experiments

Red: slow SI, fast semantic; Green: both fast

Dual Encoding

- H&S, 2009
- H&S, 2011
- Panizza, Huang, Chierchia & Snedeker (2009)
- Huang, Hahn & Snedeker
- Degen & Tanenhaus
- Hartshorne et al

Single Encoding

- Grodner et al., 2010
- Breheny, Ferguson & Katsos, (2012)
- Breheny, Ferguson & Katsos (2013)
- Huang, Hahn & Snedeker
- Degen & Tanenhaus
- Hartshorne et al.
- Huang (*most*, *start*, pc)

Underlined studies manipulated encoding

Alternative proposal

- Including numbers makes “some” less natural
  - By what mechanism does naturalness influence processing?
  - Depending on answer this may be the same account….
  - In Degen & Tanenhaus (eyetracking) naturalness doesn’t predict speed of reference resolution
  - Naturalness ratings for sentences embedded in our story task do not support (some = two)
- Bayesian proposal
  - May describe what gets computed
  - But doesn’t provide a clear story of how

Bayes Theorem

Meanings $M = \{m_0, m_1, m_2, \ldots, m_\forall\}$
Utterances $U = \{u_{\text{some}}, u_{\text{all}}, u_{\text{none}}, u_{\text{number}}\}$
QUD $Q = \{\text{qud}_{\text{all?}}, \text{qud}_{\text{any?}}\}$

$$P_{\text{listener}}(M|U, Q) \propto P_{\text{speaker}}(U|M, Q)P(M)$$

Bayes Theorem

Meanings $M = \{m_0, m_1, m_2, \ldots, m_\forall\}$
Utterances $U = \{u_{\text{some}}, u_{\text{all}}, u_{\text{none}}, u_{\text{number}}\}$
QUD $Q = \{\text{qud}_{\text{all?}}, \text{qud}_{\text{any?}}\}$

$$P_{\text{listener}}(M|U, Q) \propto P_{\text{speaker}}(U|M, Q)P(M)$$

Bayes Theorem

Meanings $M = \{m_0, m_1, m_2, \ldots, m_\forall\}$
Utterances $U = \{u_{\text{some}}, u_{\text{all}}, u_{\text{none}}, u_{\text{number}}\}$
QUD $Q = \{\text{qud}_{\text{all?}}, \text{qud}_{\text{any?}}\}$

$$P_{\text{listener}}(M|U, Q) \propto P_{\text{speaker}}(U|M, Q)P(M)$$

Awesome, how does the listener get that?

Option 1: ask 100 people on AMT?

Option 2: use stored knowledge (of Bill & Judy and the soccer balls?)

Option 3: run a production simulation (our proposal)

Outline

- Grounding assumptions
- A modest proposal
- The data behind the proposal
  1. Implicature typically takes time and effort
  2. Instant SI's occur only when pre-encoding is plausible
  3. SI proficiency develops slowly
  4. In disordered populations SI patterns with language ability

The primary observation

Children often accept under informative scalar terms in judgment tasks

- Accept “might be” in context of MUST BE (Noveck, 2001)
- Accept “started” for FINISHED (Papafragou & Musolino, 2003)

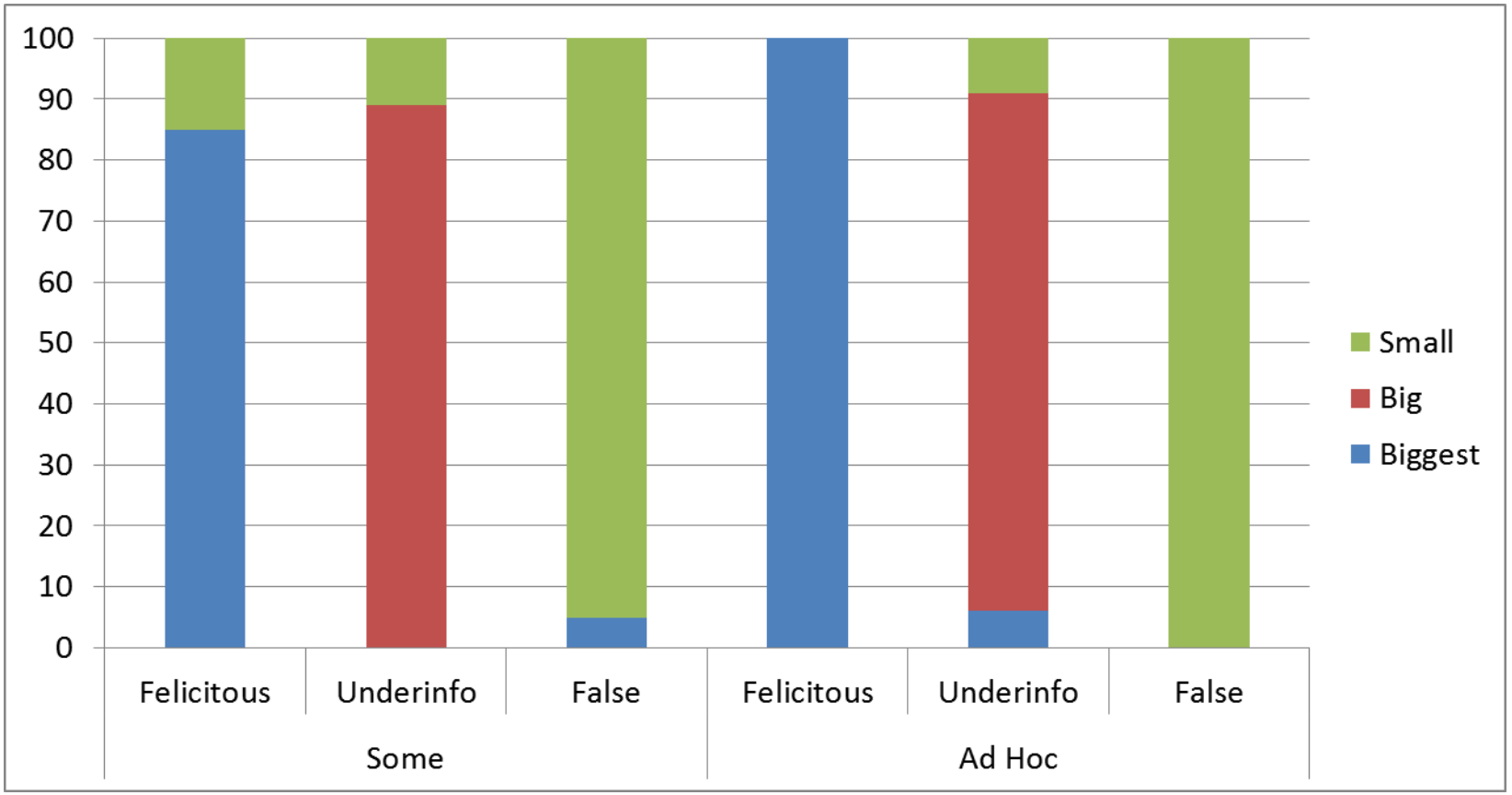
Possibility 1: Children must acquire a single discrete skill (implicature)

Non-starter: there is too much variation

- Performance heavily task dependent (Papafragou & Tantalou, 2004; Pouscoulous, Noveck, Politzer, Bastide, 2007)
- Instructions matter (Papafragou & Musolino, 2003 i.a.)
- Variation across scalar terms
- Age range success ~3-10

Possibility 2: Children are simply tolerant
(Katsos & Bishop, 2011)

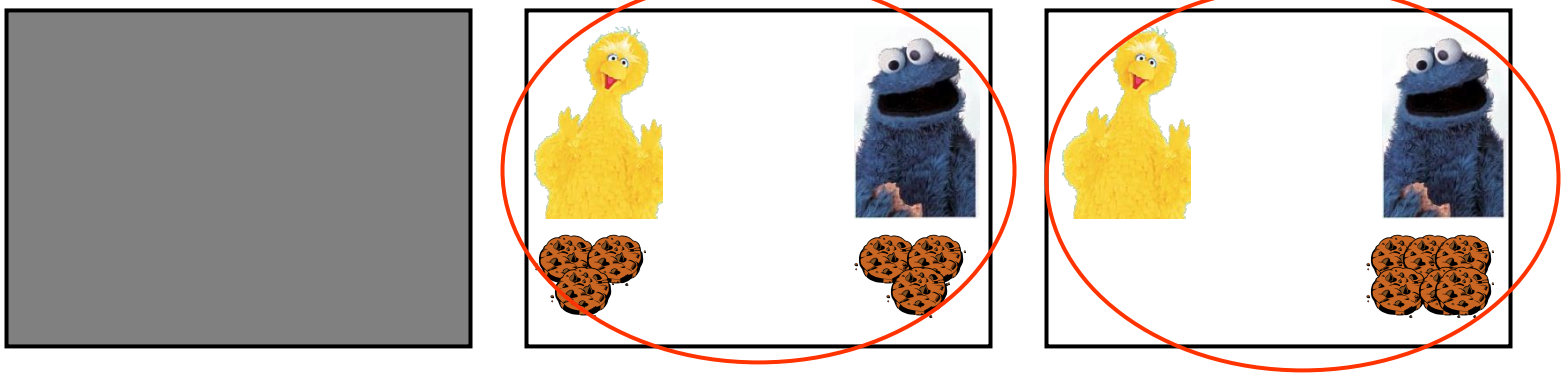
- 5 year olds succeed with 3 point scale



Tolerance can’t explain it all

- Younger children fail at selection tasks
  - Huang, Spelke, & Snedeker 2013 (2;6-4:0)
  - *“Can you give me the box where Cookie Monster has some of the cookies?”*

kids pick either one



See also Hurewitz et al., 2006

Tolerance can't explain it all

- Generic bias (Leslie & Gelman, 2012)
  - Adults and children misremember universal statements as generics (all dogs → dogs)
  - 3 yr olds *also* misremember "some" statements as generics (some dogs → dogs)
  - Suggests they aren't generating implicature
- Processing failure (Huang & Snedeker, 2009, Dev Psych)
  - Adults slower to interpret underinformative *some* than felicitous *some*
  - Children are not!

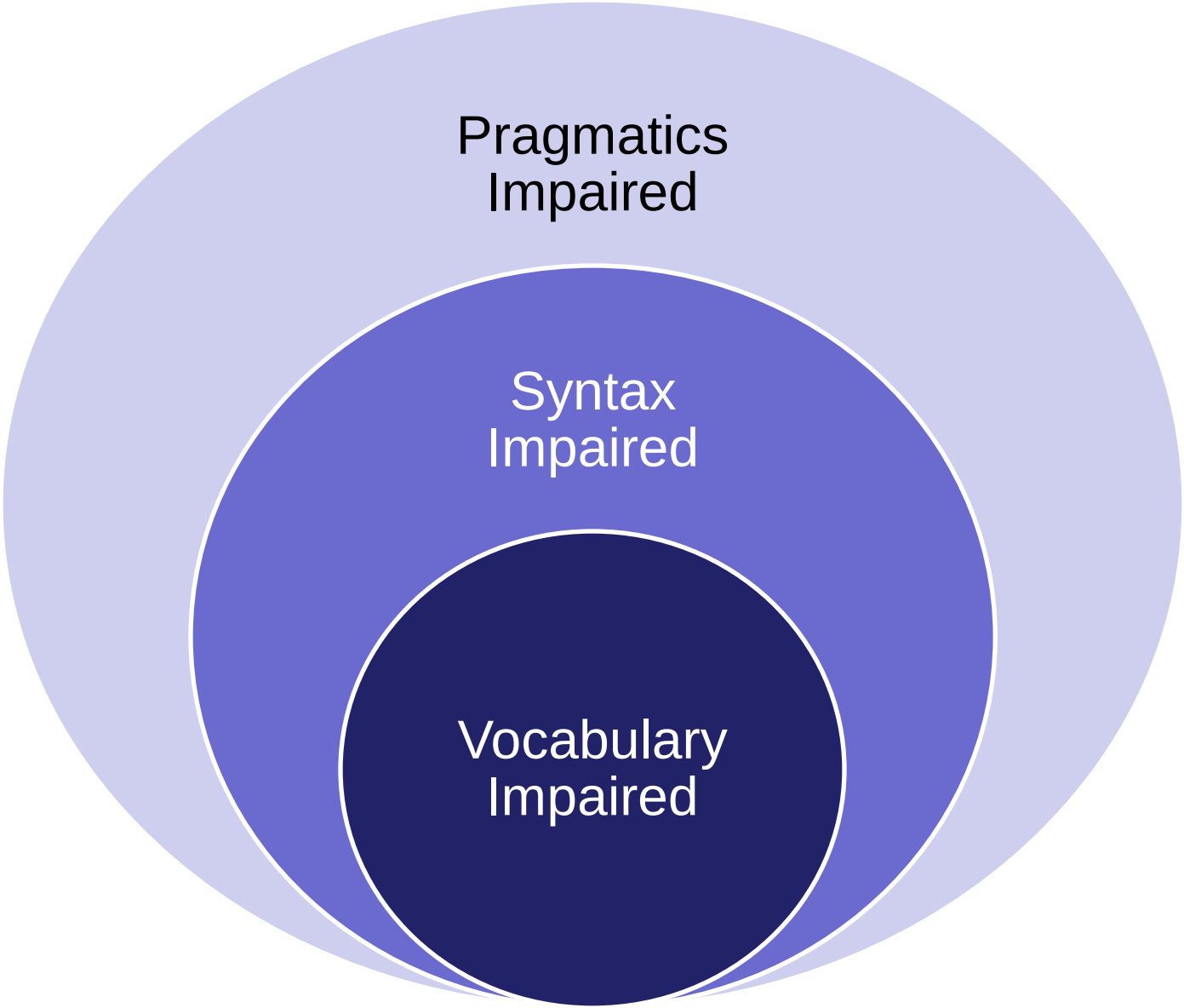
Possibility 3: processing account

- Computing SI without pre-encoding is effortful
  - See above
  - Children fail to pre-encode in contexts where adults do (Huang, data)
- Children have difficulty retrieving scales (Barner, Brooks & Bale, 2012)
- Children have difficulty using top-down cues (Snedeker, 2013)
  - SI may involve generating higher-level information to enrich interpretation
  - Such loops unfold over time (see Dell, 1986)
  - Slower processing = fewer time steps….
- As they become faster more efficient processors, they may be able to calculate SI's more often

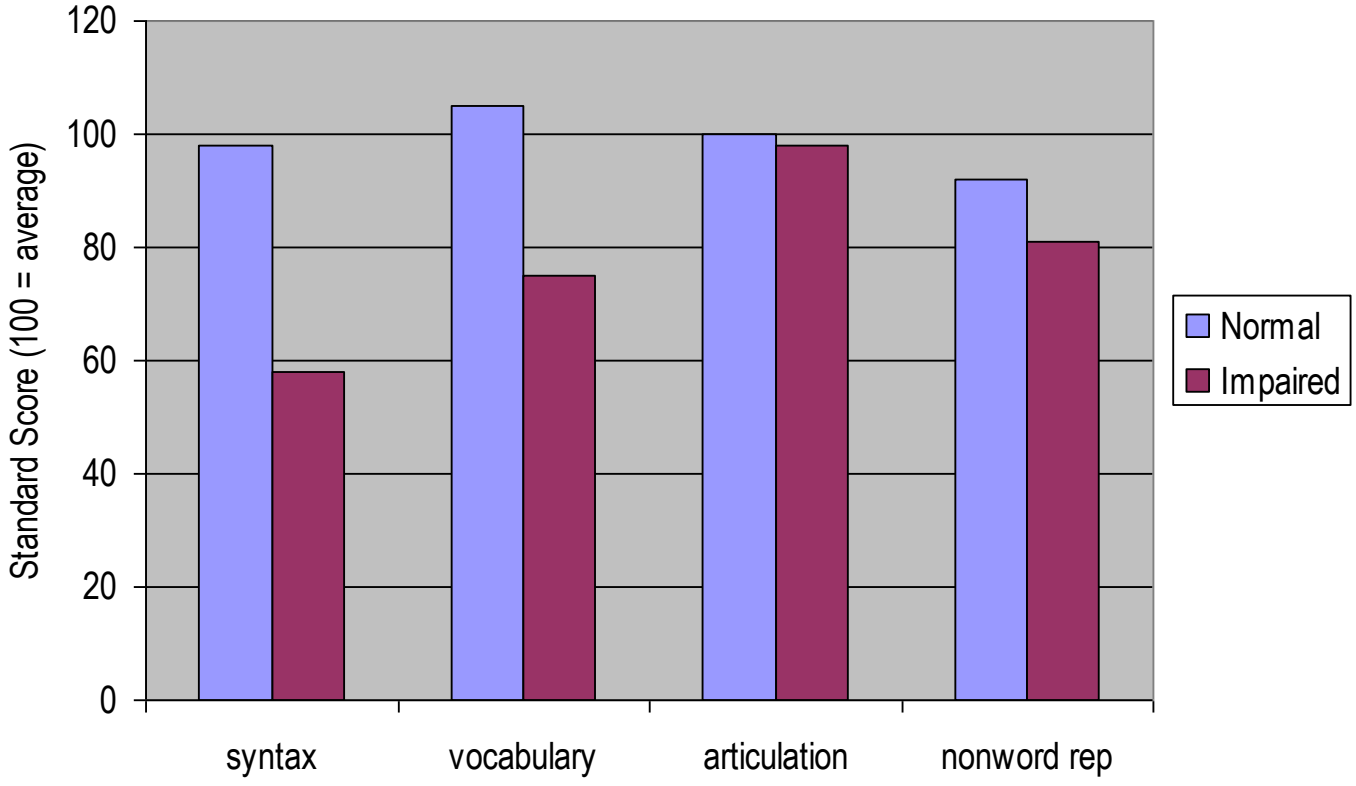
Outline

- Grounding assumptions
- A modest proposal
- The data behind the proposal
  1. Implicature typically takes time and effort
  2. Instant SI's occur only when pre-encoding is plausible
  3. SI proficiency develops slowly
  4. In disordered populations SI patterns with language ability

Communicative deficits in autism



Autism with, and without, language impairment



Kjelgaard & Tager-Flusberg (2001)

Autism and scalar implicature

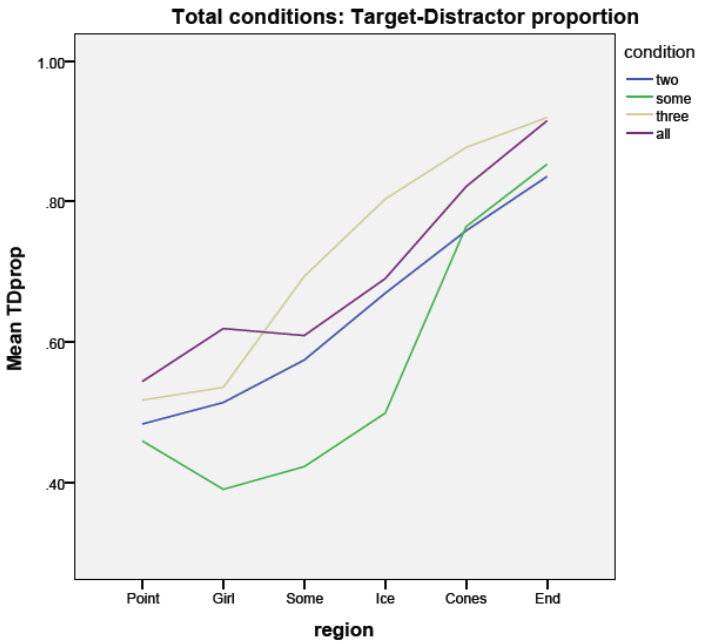
- Adults and teens with autism make SI's as often as language-matched controls (Pijnaker et al., 2008; Chevallier et al., 2010).
- Early deficit could disappear by 13
  - Deficits in Theory of Mind task only present until verbal mental age of 6-7 (Happe, 1995)
  - SI improves from 4 to 10 years
- Do persons with autism use the same process?

Our study

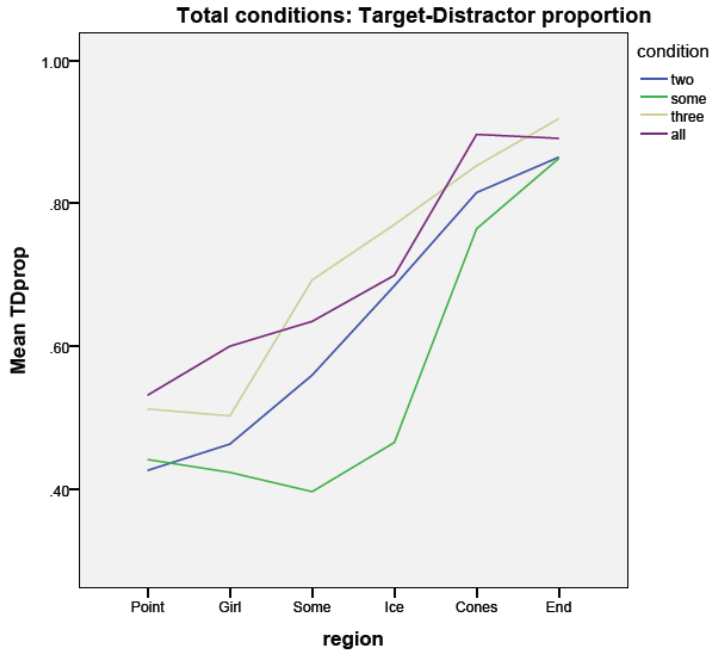
(Hahn, Huang & Snedeker, in prep)

- Goals
  - Assess likelihood of calculating scalar implicature at an age where it is rapidly changing (box task)
  - Determine whether mechanisms of comprehension are similar (visual world task)
- 6-9 year olds children
  - 40 with High Functioning Autism
  - 40 Typically Developing
  - Matched on: age, gender, CELF syntax scores

Same online processing profile



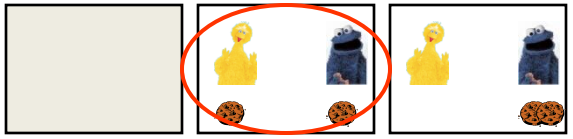
Typically Developing



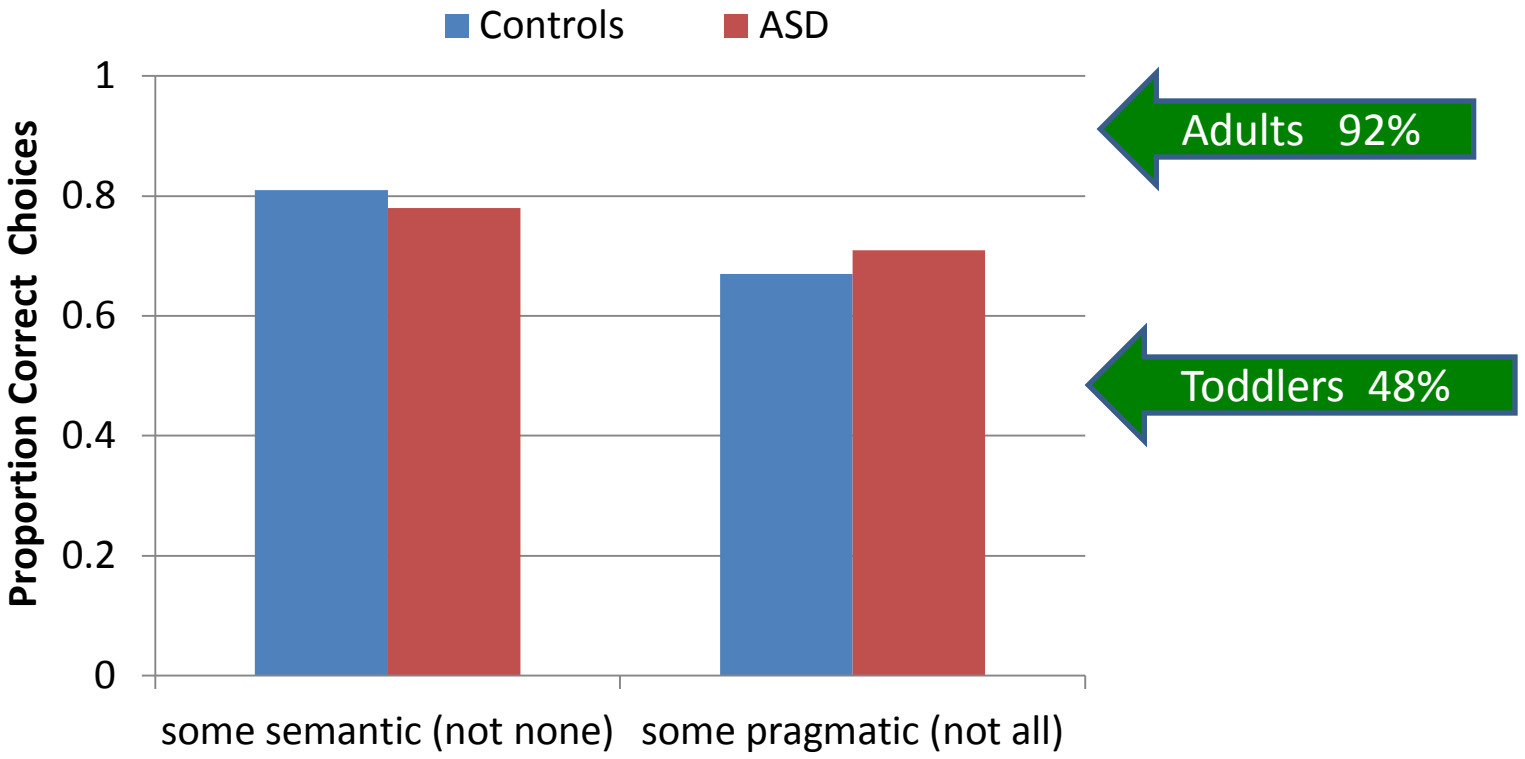
Highly Verbal ASD

Some -------

Hahn, Huang & Snedeker, in prep

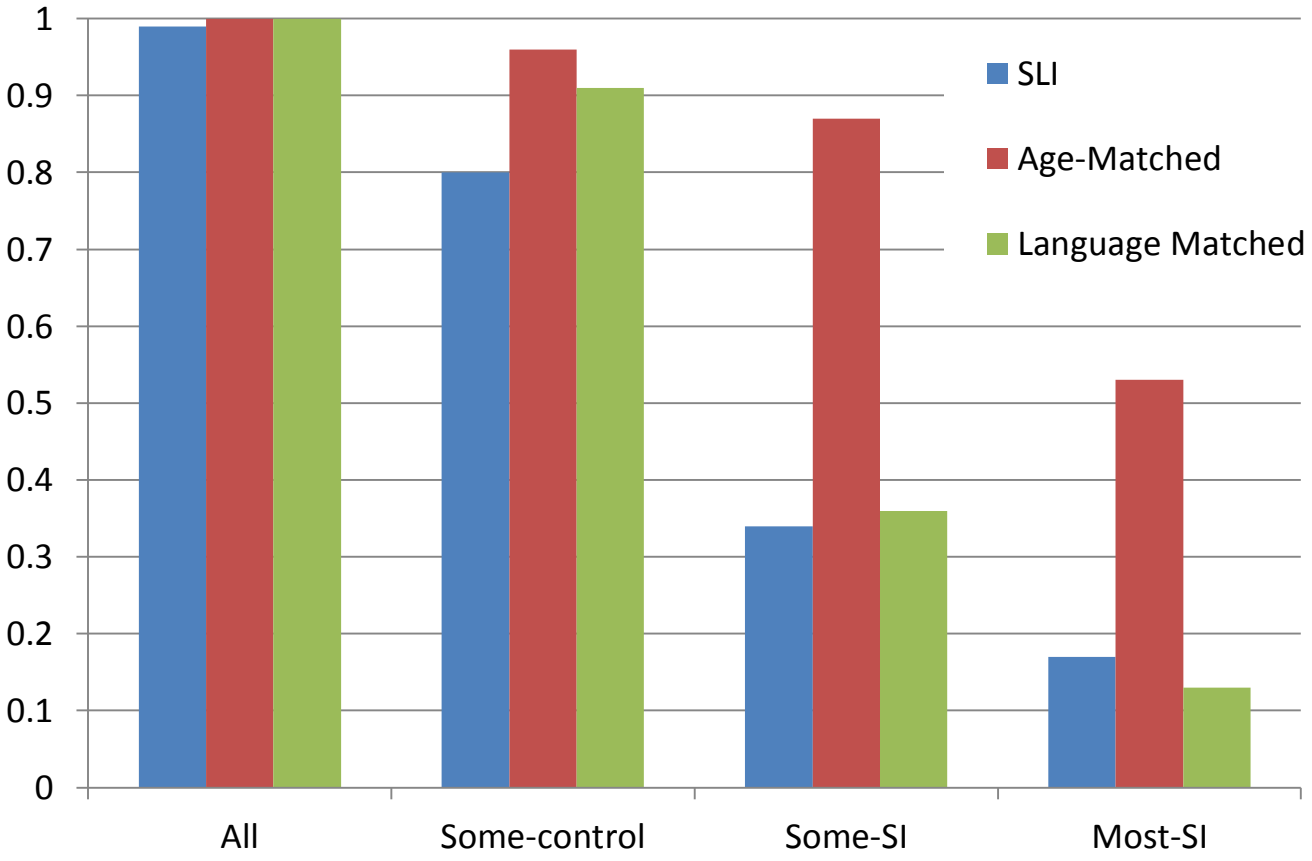


During the period where SI is developing children with ASD perform as well as controls



SI is linked to emerging language skills

Katsos, Roqueta, Clemente & Cummins (2011)



The only evidence that SI is linked to ASD….

- Nieuwland, Dittman & Kuperberg (2010)
  - "Some people have lungs/pets"
  - N400 at *pets*
  - Correlates with AQ communication scale (not social scale)

- My suspicion:
  - In college students, communication scale may capture differences in language skills

In sum

1. Implicature takes some work (bottom up)
2. But the work can be done ahead of time
   - When the conceptual encoding for each message is unambiguous
   - Listener as speaker
3. Thus SI proficiency develops gradually as children become more effective processors
4. Thus SI breaks down with language skills
   - Consistent with a distinction btw grammatical/social inferences or explicatures/implicatures?

Thank you!

- National Science Foundation & Simons Foundation
- Collaborators: Yi Ting Huang x 10, Gennaro Chierchia, Daniele Panizza, Joshua Hartshorne, Manizeh Khan & Noemi Hahn
- Assistance from: Amanda Worek, Carlyn Friedberg, Carissa Shafto, and dozens of interns