

Methods in the acquisition of implicature

Napoleon Katsos

Department of Theoretical and Applied Linguistics

University of Cambridge

Acquisition of quantity implicature

Today: Methods in the acquisition of quantity implicature

Thursday: Quantity implicature and its relatives

Friday: Semantics and Pragmatics across populations and languages

The Gricean account

(Grice, 1975; also Horn, 1984, 1992 i.a.)

- Interlocutors are expected to be co-operative and obey the maxims of conversation. The first sub-maxim of Quantity: “make your contribution as informative as is required for the purposes of the conversation.” (Grice, 1989: 25)

John: Peter is meeting a woman this evening

+> The person Peter is meeting is not his wife

Under certain assumptions: cooperation, knowledge, relevance

Other Quantity Implicatures

- (1) a. John: Did all your students pass the test?
b. Mary: Some of the students failed.
+> Not all the students failed

- (2) a. John: Why is the teacher upset?
b. Mary: Some of the students failed.

Scales and implicatures

- <or, and>
 - <some, many, most, all>
 - <one, two, ..., n>
 - <like, love, adore>
- Asymmetric entailment relations.
Terms on the right of the scale are more informative than terms on the left, counted in number of entailments:

All Fs are G \Rightarrow Most/ Many /Some Fs are G

Some Fs are G \nRightarrow All / Most / Many Fs are G

Generalised vs Particularised Quantity Implicatures

- John: Did all your students pass the test?
- Peter: Some of my students failed.
 - Not all of Peter's students failed.
- John: Did you meet Linda's parents already?
- Peter: I met her dad.
 - Peter didn't meet Linda's mom.

- Some of my students passed.
- Implies: 'not all the students'
- Scale: <some, all>
- Other generalised scales:
 - <possible, certain>
 - <or, and>
 - <like, love>
 - ...
- For each term of a **generalised** scale, we know independently of context the other terms of the scale
- I met Linda's dad.
- Implies: 'but not her mom'
- Scale: <{dad}, {mom}, {dad and mom}>
- BUT ONLY IN SPECIFIC CONTEXTS
- For terms of **ad hoc** scales, we cannot know independently of context the terms of the scale

(1) The Bride: Have you seen Tommy?
Bill: Big guy in the tux?
The Bride: Yes.
Bill: Then I saw him. I like his hair.
The Bride: You promised you'd be nice!

(Kill Bill, Volume 2; script by Quentin Tarantino;
Uma Thurman as the Bride; David Carradine as Bill)

The particularised – generalised conversational implicature distinction (PCIs – GCIs)

- “Some implicatures are induced **only** in a special context..., while others go through **unless** a special context is present.”
- See Horn’s example from ‘When Harry Met Sally’ (Horn, 2006)
- Putting aside this distinction, for Grice, Horn and other traditional neo-Gricean approaches, both PCIs and GCIs are conversational implicatures and should exhibit the typical properties of implicature-hood

Alternative accounts

- For a specific subset of implicatures, scalars
 - Levinson (2000)
 - Chierchia (2004)
 - Chierchia (2006), Chierchia, Fox & Spector (2012), Chemla & Singh (2014)

Overview

- Pragmatic view
 - Level: speaker intentions
 - Form: a pragmatic mechanism, with disjunctive syllogism (A or B, not B, therefore A) operating on speaker's intentions
- Non-pragmatic views
 - Level: grammar (syntax or semantics) or default pragmatics
 - Form: 2-system approach:
 - A non-pragmatic mechanism derives an interpretation which an implicature
 - A pragmatic (context-tracking system) either to license or to suspend the application of the non-pragmatic mechanism

A big question for early Xprag folk

- What are the mechanisms involved in implicature derivation?
- A wealth of (possibly theory-critical) experimental findings
 - A developmental story
 - A processing story
 - Methodological insights

- Richard: Somebody's going to have an ice-cream today!
- Ned (24 months): No! Ned is!

- Noveck, 2001: Experiment in French “certain”/ “tous”
- 3 age groups:
 - Group 1: 8 year old (n=31)
 - Group 2: 10 year old (n=30)
 - Group 3: adults (n=15)

<u>Sentence Type</u>	<u>Correct Response</u>	<u>Age</u>		
		7-8 (n=31)	10-11 (n=30)	Adults (n=15)
<u>Utterances expressed with all</u>				
<u>Absurd (false)</u> (e.g. All chairs tell time)	No	93%	99%	99%
<u>Appropriate (true)</u> (e.g. All elephants have trunks)	Yes	91%	99%	96%
<u>Inappropriate (false)</u> (e.g. All dogs have spots)	No	86%	99%	96%
<u>Utterances expressed with some</u>				
<u>Absurd (false)</u> (e.g. Some stores are made of bubbles)	No	95%	99%	98%
<u>Appropriate (true)</u> (e.g. Some birds live in cages)	Yes	84%	90%	99%
<u>Inappropriate (true though pragmatically infelicitous)</u> (e.g. Some giraffes have long necks)	Yes	89%	85%	41%

Some giraffes have long necks...

- This task is relatively difficult.
- Although particular ages are noted Noveck (2001) carefully avoided making the claim that the general capacity to infer implicatures is linked to a particular chronological age.

Papafragou & Musolino, 2003: task specific and training effects

- ◆ Quantifier scale: < *oli, meriki* > ('<all, some>')
- ◆ Aspectual scale: < *teliono, arxizo* > ('<finish, start>')
- ◆ Numerals: < *trīs, duo* > ('<three, two>')

Act out scenario: some/all



The end of the story



The end of the story: two/three

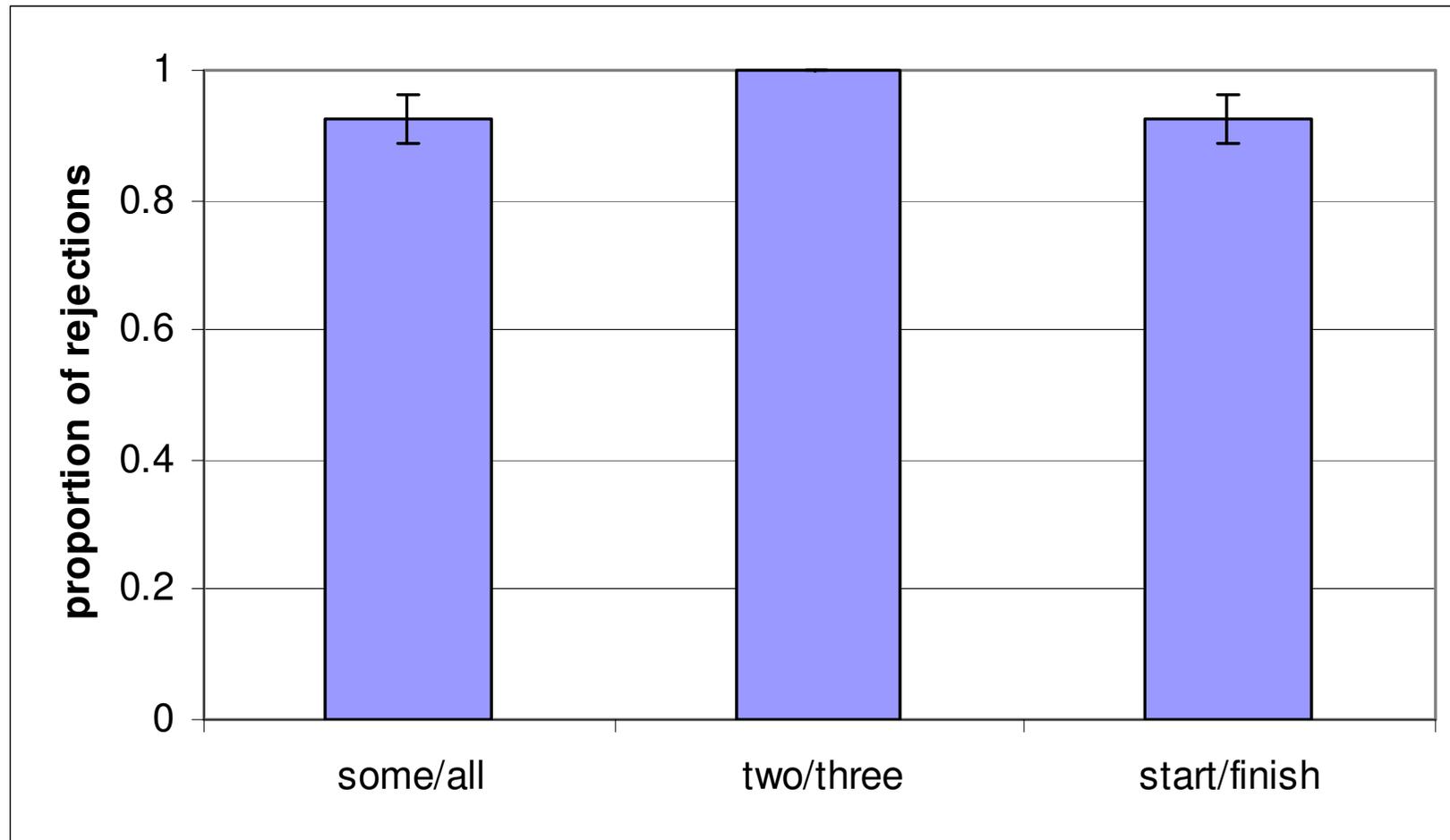


Act out scenario: start/finish



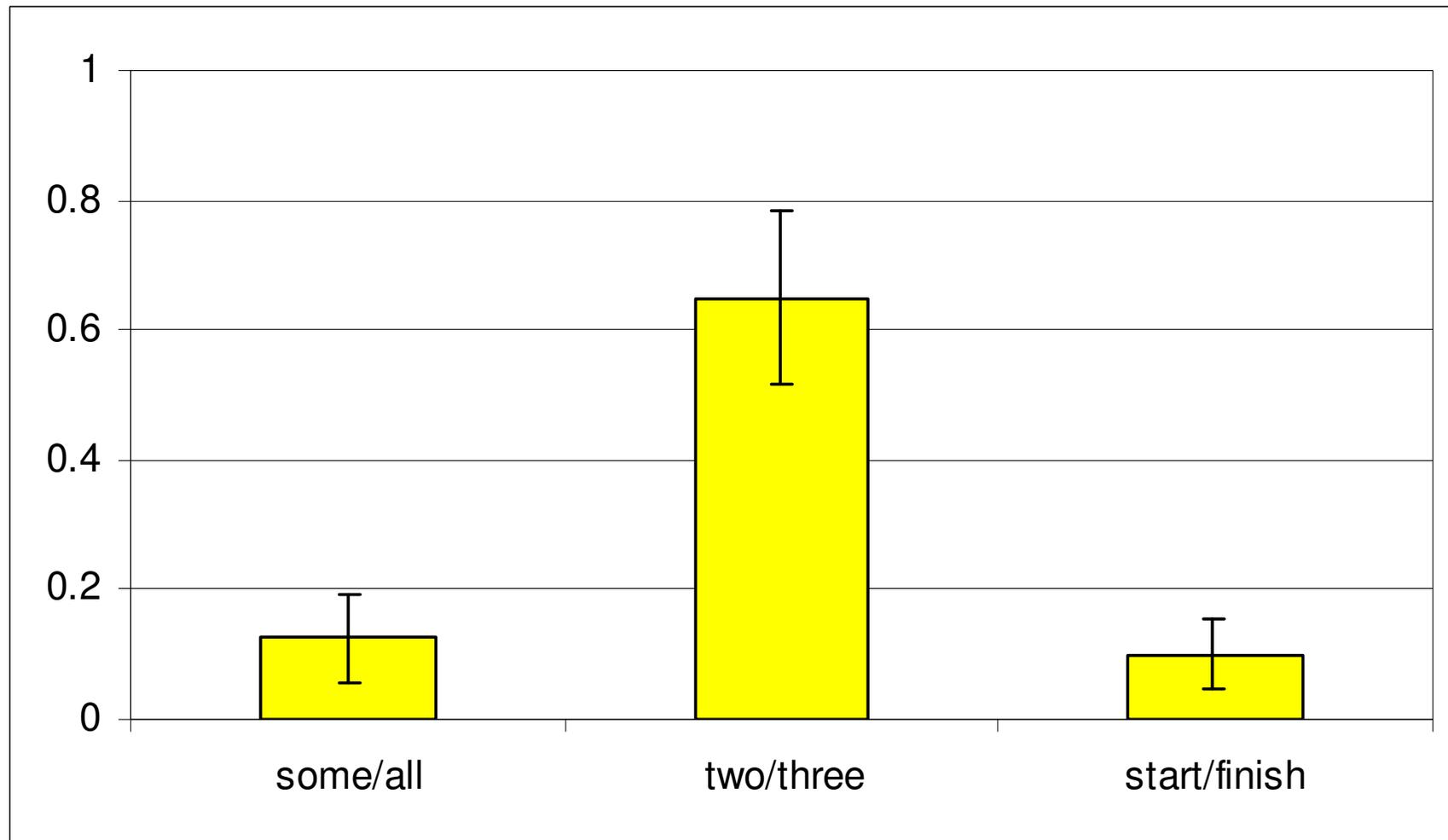
Results: adults

Proportion of 'informative' responses, i.e. with the SI



Results: children

Proportion of 'informative' responses, i.e. with the SI



Papafragou & Musolino 2003: task specific and scale specific effects

- Maybe children are not given enough motivation to explore the full meaning of an utterance
- Hypothesis: Children \neq adults: not a difference in competence but a difference in the ability to detect the goal of the interaction, the appropriate expectations of informativeness.

One of the warm-ups: “A small animal with fluffy ears”

“Did Minnie say it well? Can we say it better?”



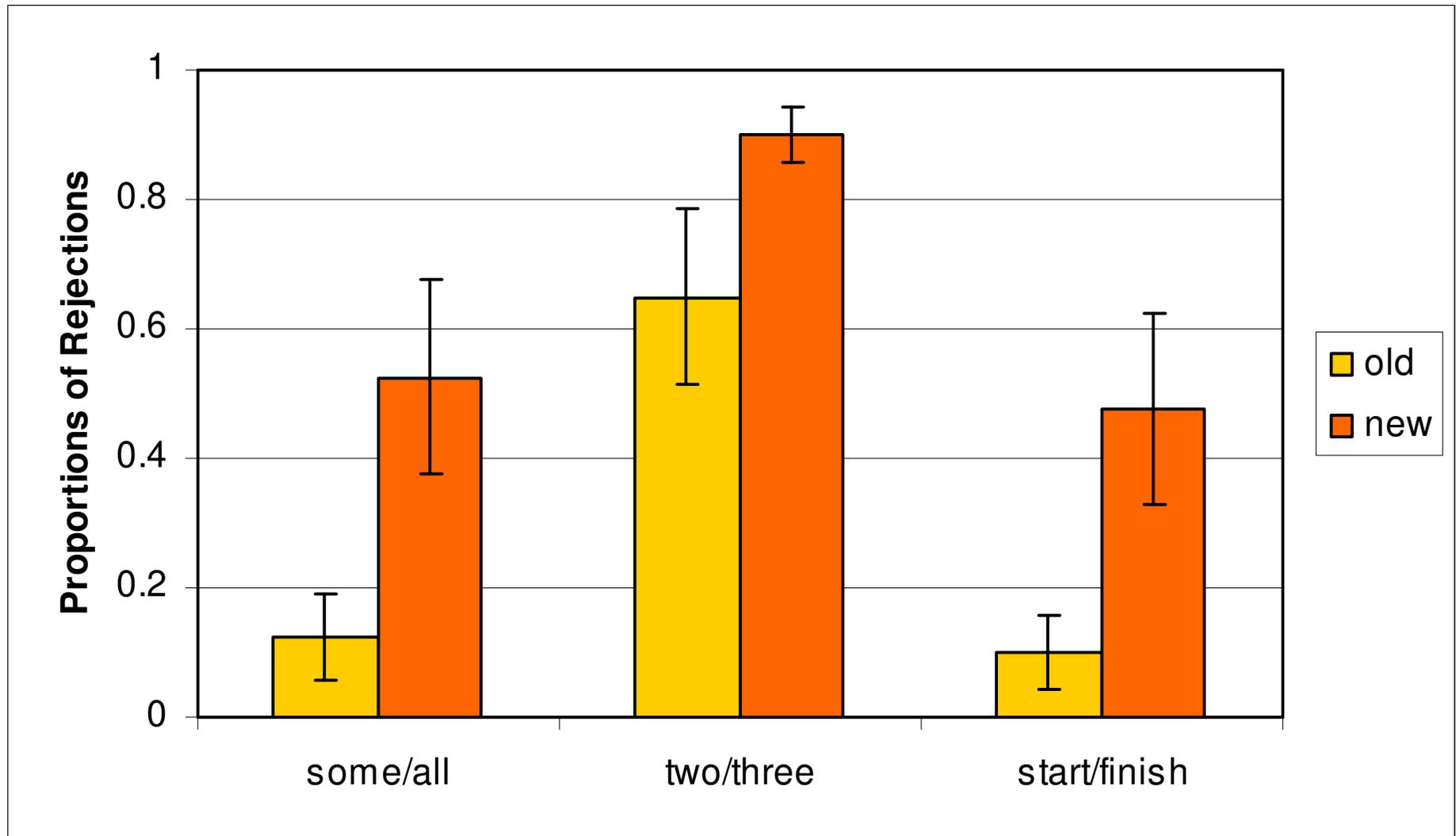
One of the stories: some/all



The end of the story



Results: P&M 2003 Exp 1 vs. Exp 2



Interim conclusions: sentence judgement tasks

- Noveck 2001: 8 & 10-year-olds do not generate SIs and they accept under-informative sentences
 - They don't have the processing resources to engage in the processes that derive implicatures (informed by resources/cognitive effects trade-off in Relevance Theory). See also Poscoulous et al, 2007
- Papafragou & Musolino: 5-year-olds do well with some implicatures
 - Children become much better when they are trained to detect pragmatic infelicity, perhaps they do not have adult-like expectations of what counts as 'informative enough' in a situation

Sentence-evaluation tasks

- A productive paradigm for child studies
 - Guasti et al (2005)
 - Foppolo et al (2012)
 - Barner et al (2011)
 - Skordos & Papafragou (2016)

Latest News 

The presence of alternatives, <all>, raises the rates of implicature, especially if they are perceived as being relevant to the Question Under Discussion

- Taking stock
 - Jury is out for theory-critical conclusions
 - But why are 5-year-olds not reaching adult-like rates?

Pouscoulous, Noveck, Politzer and Bastide (2007)

Can children generate SIs without 'training'? (a la P&M)

YES: if one manipulates the task demands

Experiment 1: replicated Noveck (2001)

Some turtles are in the boxes

In a scenario where all the turtles were in the boxes.

Results: Similar to Noveck (2001): 91% of the children judged the underinformative utterance to be true (No SI).

Experiment 2

The task was made less complex by making the following changes :

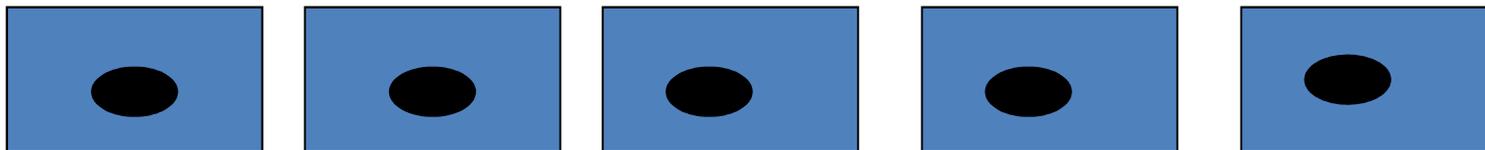
- Simpler presentation concerned only tokens and boxes
- Participants performed an action rather than making a truth-judgement

Participants

147 children, 4 to 8 years old

21 adults

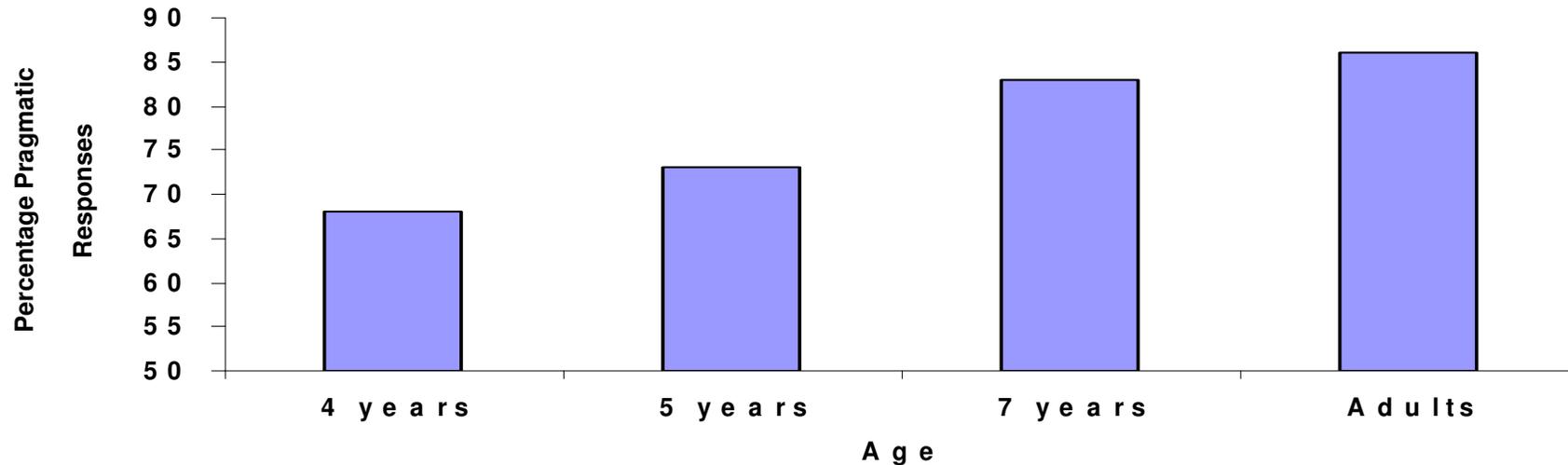
CRITICAL TRIAL



Action based evaluation

I would like some of the boxes to have a token

Pouscoulous et al; Experiment 2 – Results



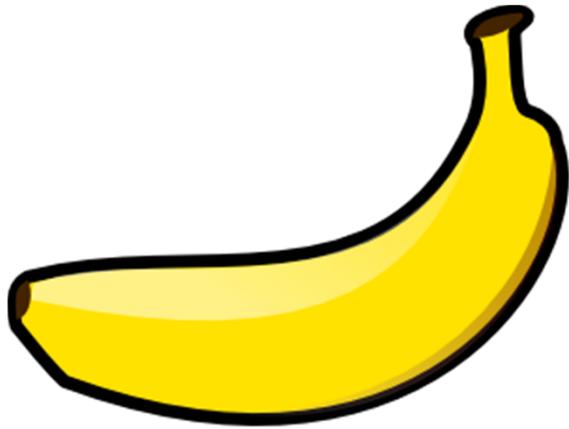
- A high rate of informative responses for all participants,
- **even the four-year-olds** (25 out of 34) are significantly above chance
- Informative** responses increase with age

Pouscoulous et al Experiment 2 - conclusions

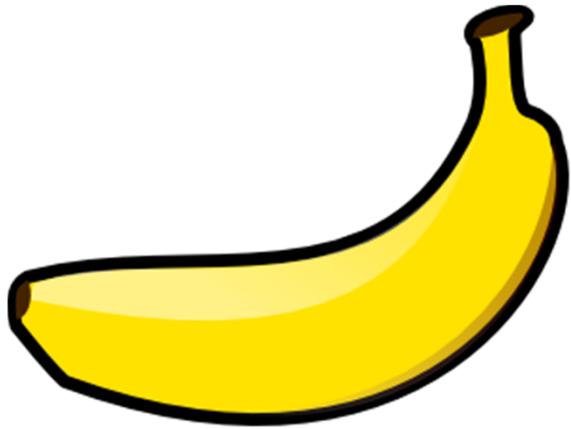
- Young children are able to spontaneously make scalar implicatures when performing a simple enough task.
- But still not-adult like



Show me the blicket



Show me one



Word-learning by exclusion

	Control Condition	Critical Condition
Mean selection of target object (novel) out of 6 pairs	3.30	4.90
Significance (compared to chance 3/6)	$t(8) = 0.64, p > .05$	$t(8) = 3.94, p < .01$

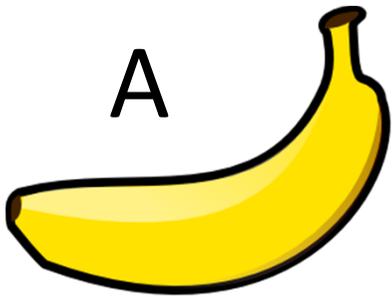
Markman & Wachtel (1988) Experiment 1,
3-4 year olds, extended to 2-year olds

How can this be accounted for?

Mutual Exclusivity

“leads children to expect each object to have only one label”

(Markman, 1990:57)



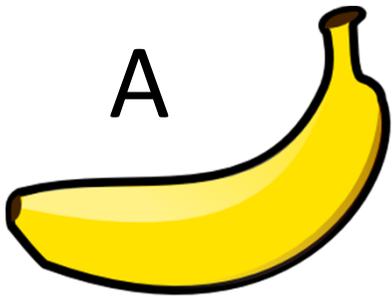
Mutual Exclusivity

(Markman & Wachtel, 1988)



1. 'Blicket' could be A or B
2. A already has a label, and each object only has one label
3. Therefore it cannot be A
4. Therefore B

Disjunctive
syllogism



Mutual Exclusivity

(Markman & Wachtel, 1988)



1. 'Blicket' could be A or B
2. If 'blicket' is A, A does not have a label
3. A does have a label,
therefore 'blicket' is not A
4. Therefore B

Modus
tollens

Disjunctive
syllogism

Conventionality & Contrast

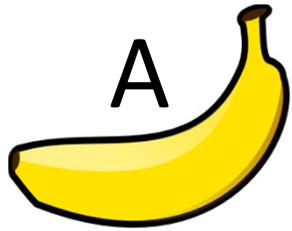
- **Conventionality**

for certain meanings, speakers assume that there is a conventional form that should be used in the language community

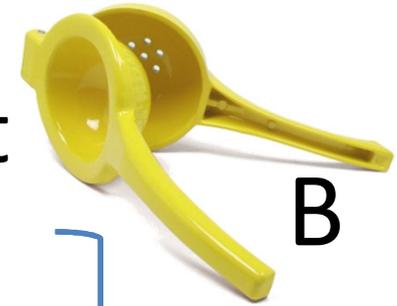
- **Contrast**

speakers assume that any difference in form signals a difference in meaning

(Clark, 2003:133)



Conventionality & Contrast



1. The speaker could be intending to refer to A or B
 2. If they meant A, they would have said 'banana'
 3. They said 'blicket', **so they do not mean A**
 4. Therefore, they intend to refer to B
- (Therefore 'blicket' refers to B)

Modus
tollens

Disjunctive
syllogism

Conventionality & Contrast and implicature

- Grice's Maxim of manner: be perspicuous
 - Avoid obscurity of expression
 - Avoid ambiguity

Obscure = non-conventional

Implicature

1. By 'some of the apples' the speaker could intend to refer to a situation with some but not all apples or to one with all apples.
2. If they meant to refer to S with all, then they would have said 'all' (because they know that 'all' would be true of that S, because 'all' is more informative, because 'all' is relevant, and assuming that they are cooperative and competent)
3. They didn't say 'all', therefore they don't mean to refer to the S with all
4. Therefore, they intend to refer to the S with some but not all

Word Learning

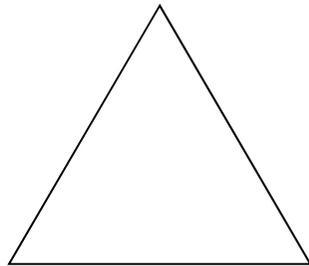
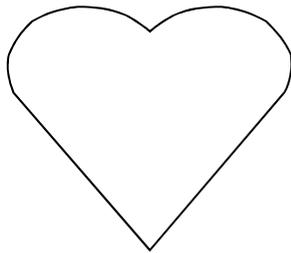
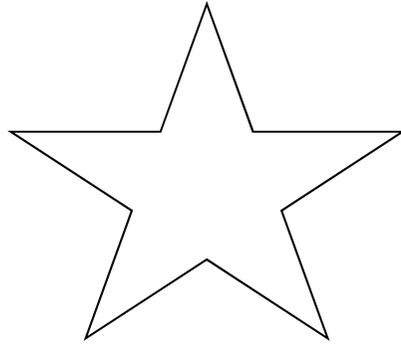
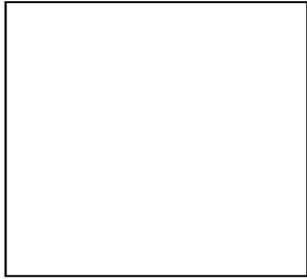
1. By 'blicket' the speaker could intend to refer to either the banana or the unknown object
2. If they meant the banana, they would have said 'banana' (because this is the conventional label and assuming they are cooperative and competent)
3. They did not say 'banana', so they don't mean to refer to the banana
4. Therefore, they intend to refer to the unknown object

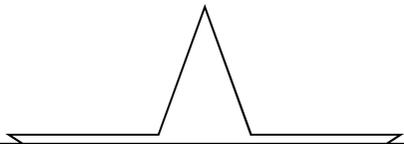


'tolerant little Grices'



Katsos & Smith, 2009
Katsos & Bishop, 2011

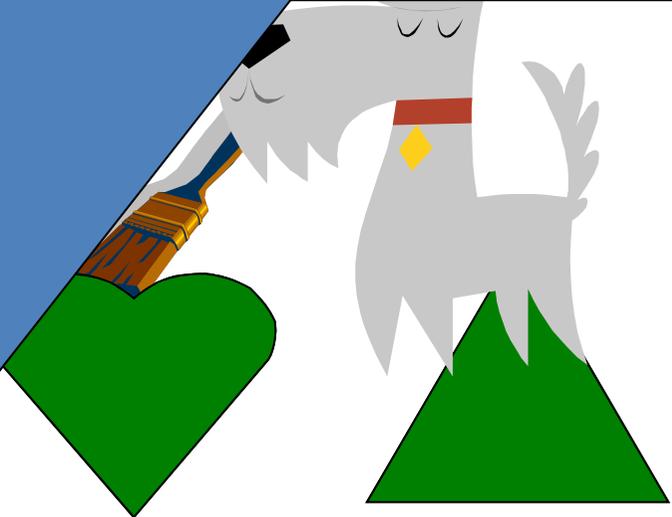


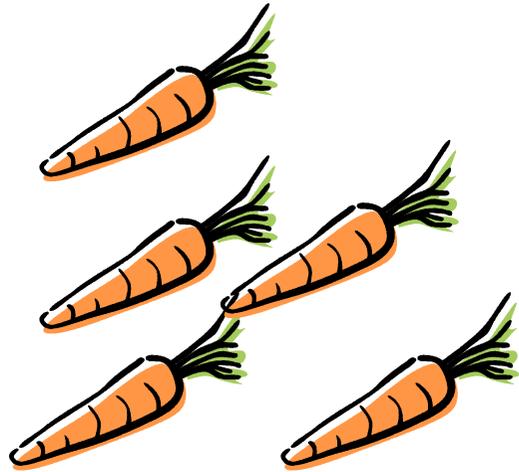


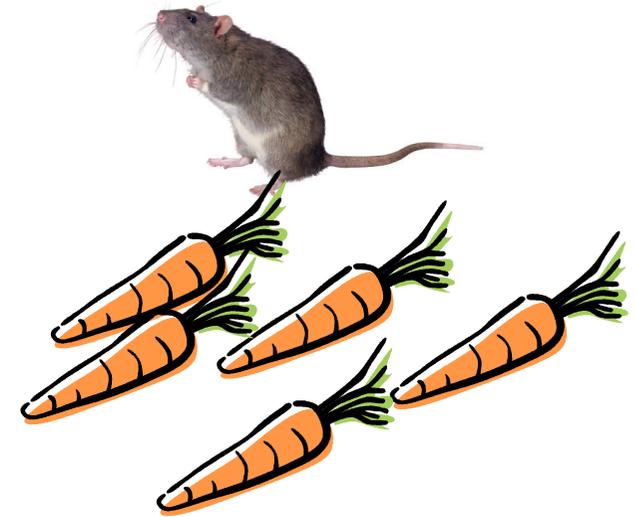
Experimenter:
"Mr Caveman, what did the dog color?"

"The dog colored the triangle."

TO PARTICIPANT: "Did he say it right?"







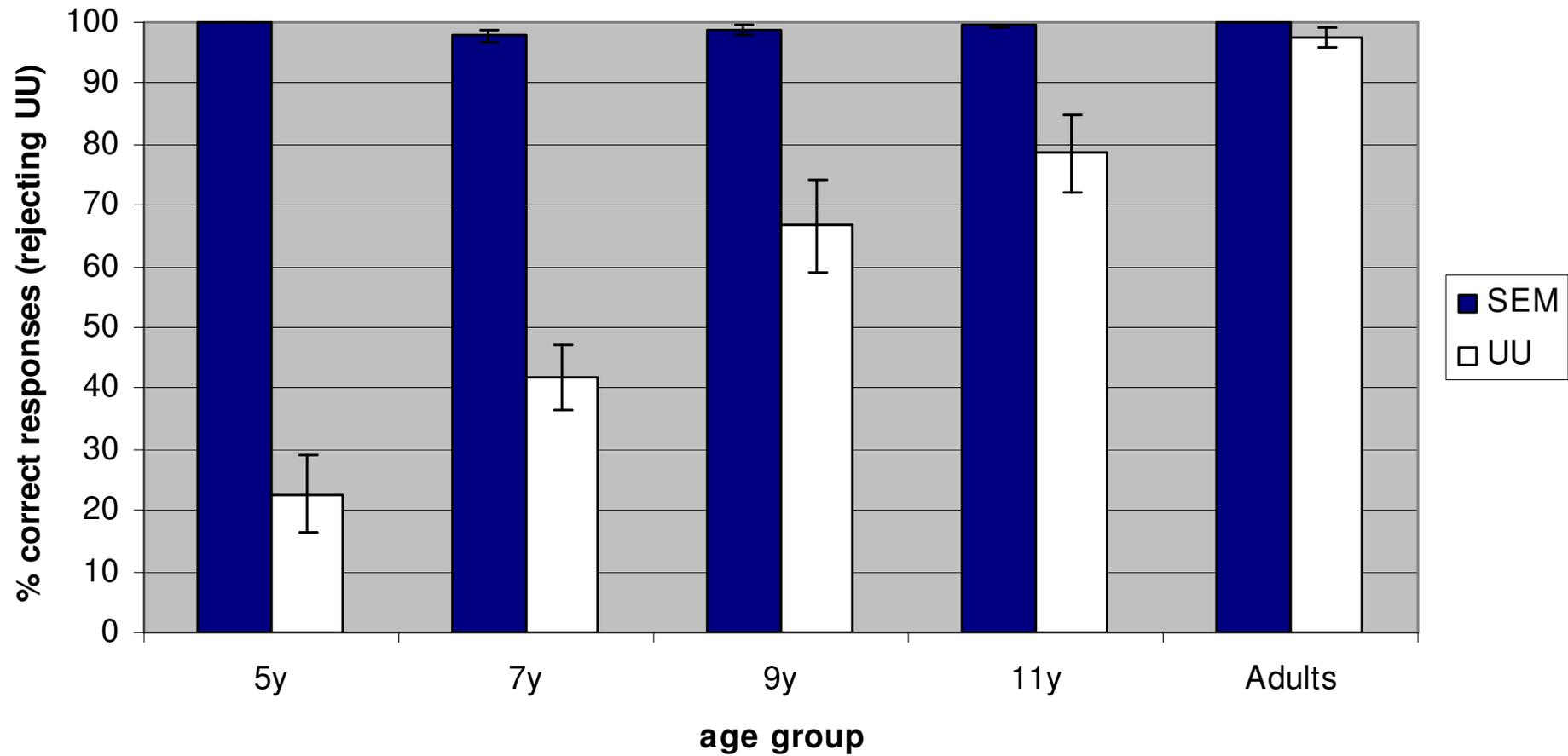
Experimenter:
“Mr Caveman, what did the mouse pick up?”

“The mouse picked up some of the carrots”

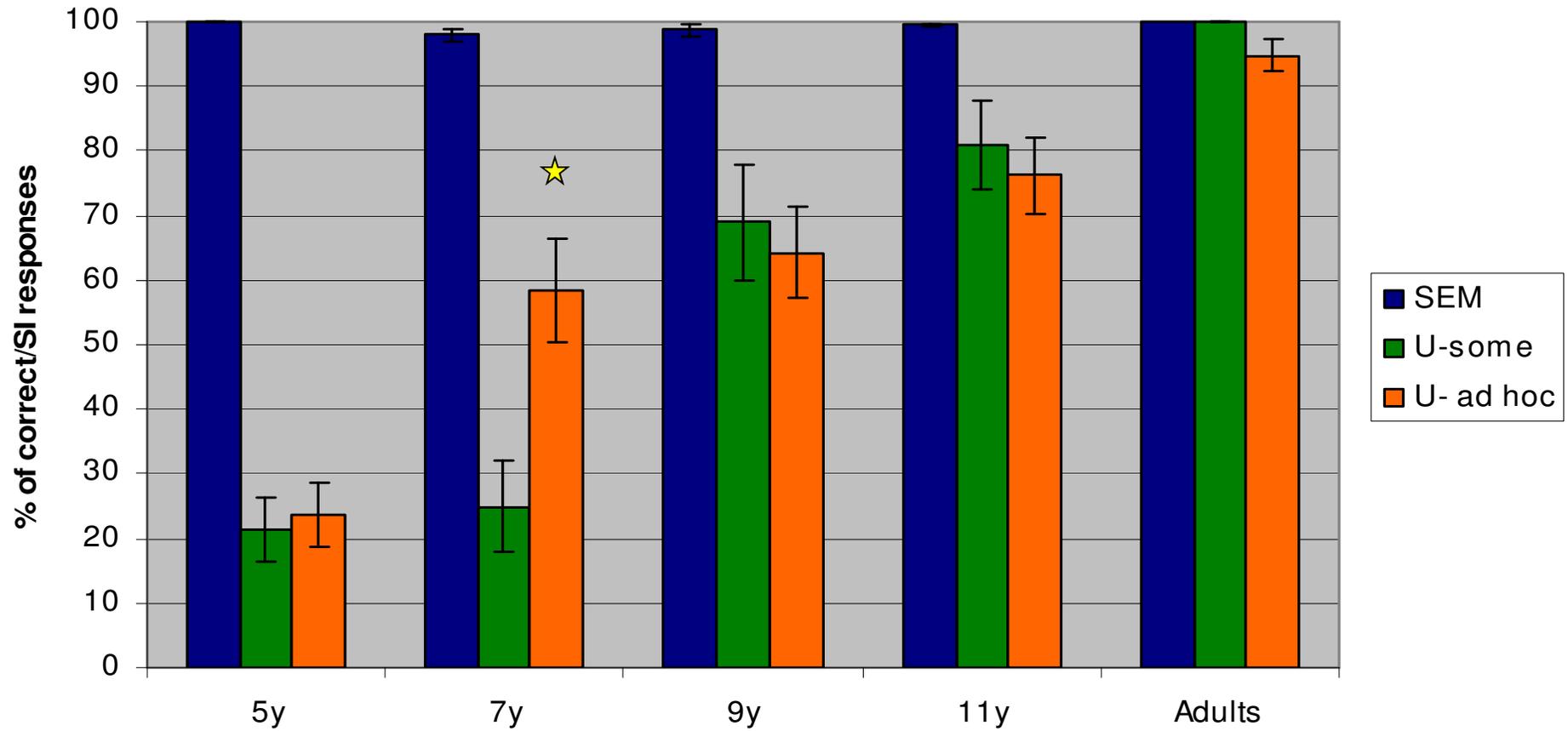
TO PARTICIPANT: “Did he say it right?”



Mr Caveman - Description mode



Mr Caveman - Description - Some vs Ad hocs



Something interesting with regards to the adults...

Categorical rejections: “**No**...he picked up **all** the carrots”
“**No**...he painted **the heart too**”

Revisions and hedges:

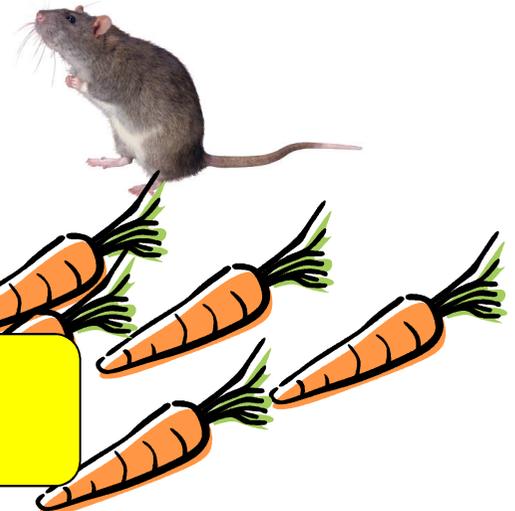
“**Yes**, ... all the carrots / **Yes**, ...**but** the heart as well”

“He was right...**but not completely**”

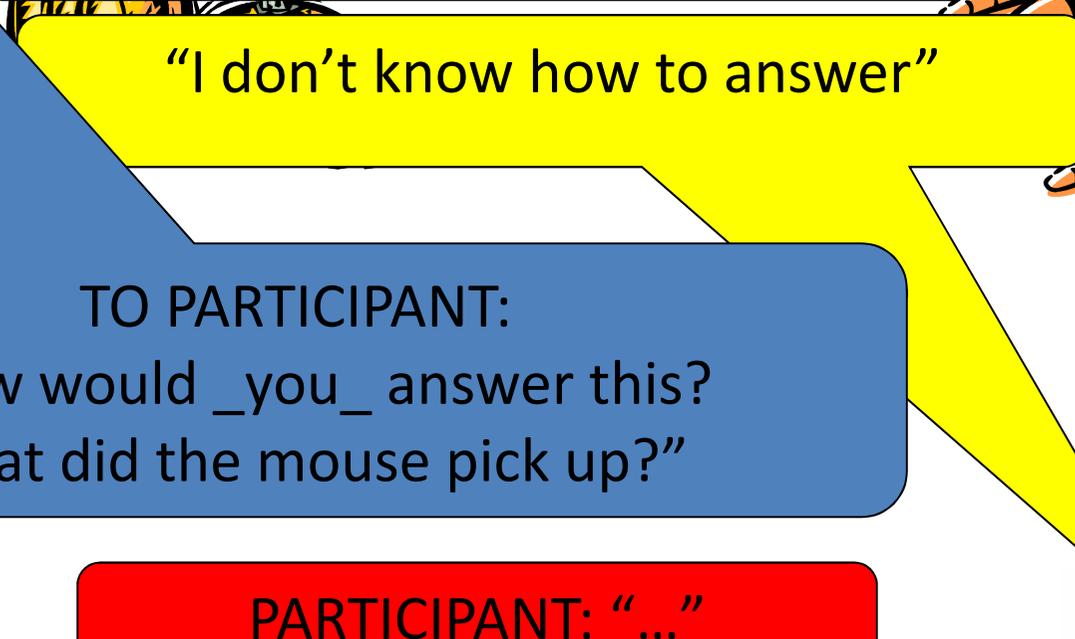
“Not wrong...not right either...” / ‘**Half** right!’ / “I can’t say”

Adults are more *lenient* towards violations of informativeness:

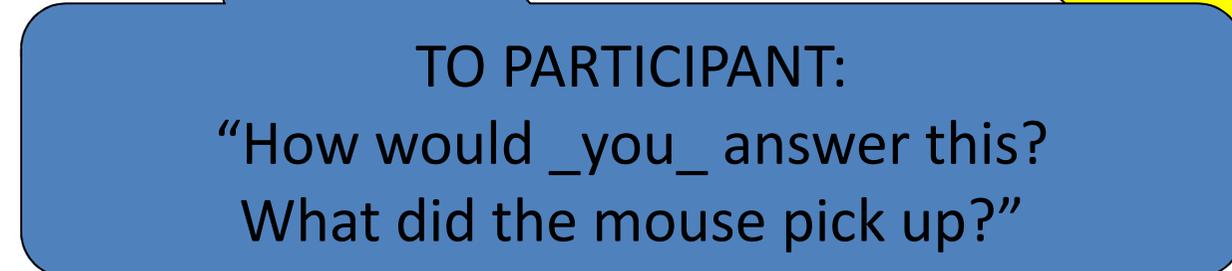
- About 30% of adult responses for U-utterances were indirect and not categorical...
- But 100% of rejections of semantically-false utterances were categorical...



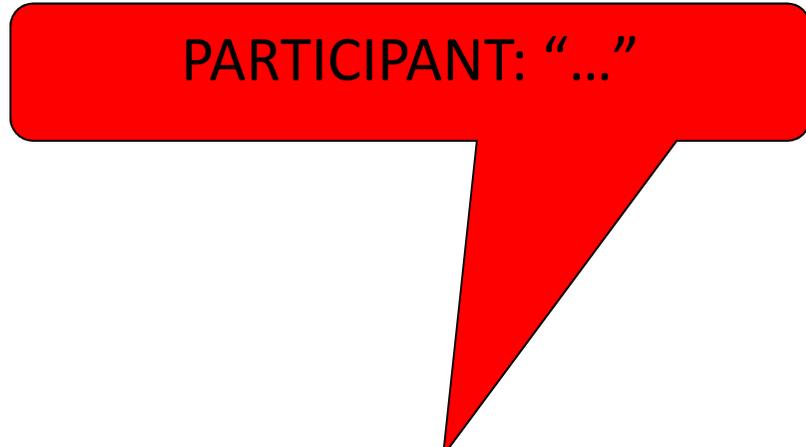
Experimenter:
“Mr Caveman, what did the mouse
pick up?”



“I don’t know how to answer”



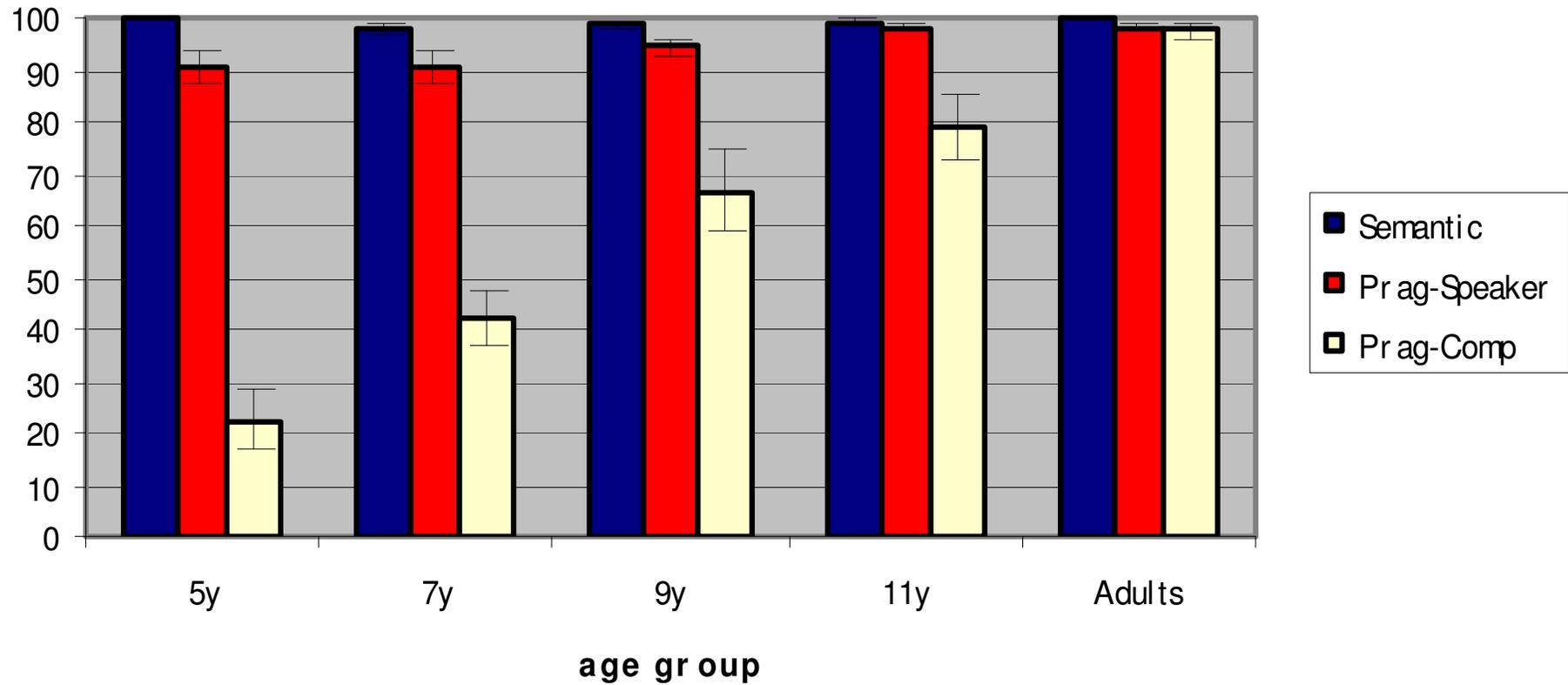
TO PARTICIPANT:
“How would you answer this?
What did the mouse pick up?”



PARTICIPANT: “...”



Pragmatic Speaker - Pragmatic Comprehender asymmetry

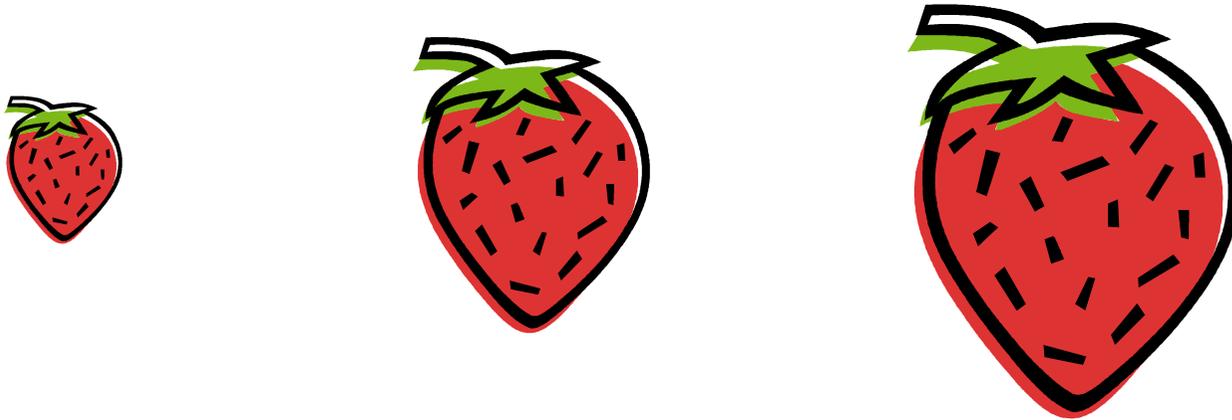


Pragmatic Tolerance

True pragmatic competence could be revealed when the pragmatically-appropriate response does not require the participant to **reject** the critical utterance

- ☛ Rating on a scale rather than binary judgment
- ☛ Sentence-Picture Matching task
- ☛ Action-based task

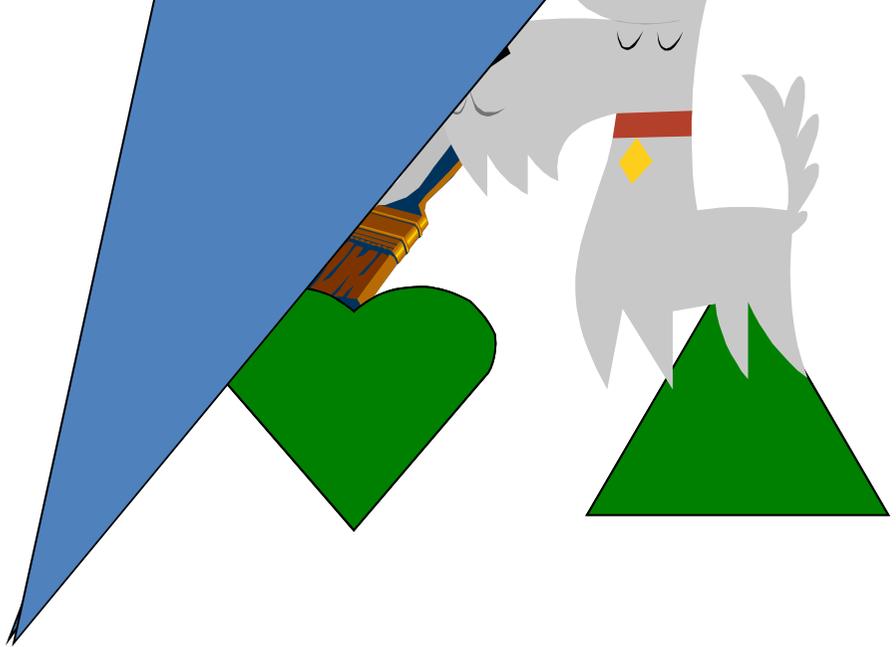
Katsos & Bishop (2011) 5 ½ year-olds and adults
'small, big, large strawberry'

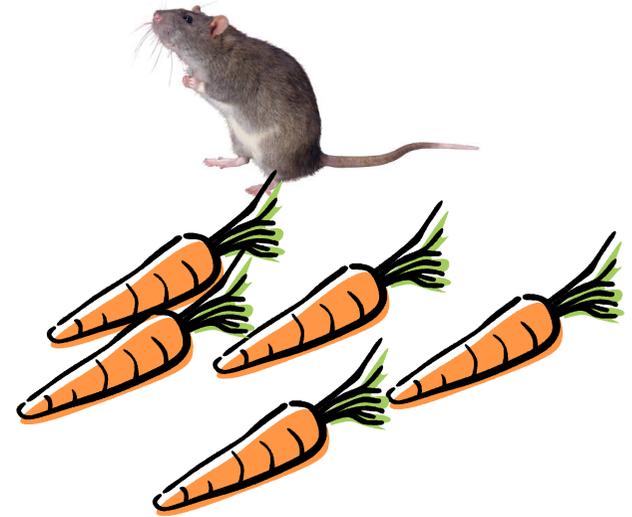


Experimentor:
"Mr Caveman, what did the dog color?"

"The dog colored the triangle."

D PARTICIPANT: "Which strawberry will you give him?"





D PARTICIPANT: "Which strawberry will you give him?"

Experimenter:
"Man, what did the mouse pick up?"

"The mouse picked up some of the carrots"

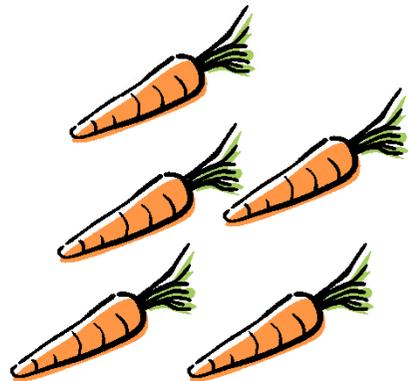


Type of utterance	Type of response	Scalar	Non-scalar	Total
Optimal	3 – ‘huge’	85	100	92.5
	2 – ‘big’	0	0	0
	1 – ‘small’	15	0	7.5
Under-informative	3 – ‘huge’	0	6	3
	2 – ‘big’	89	85	87
	1 – ‘small’	11	9	10
False	3 – ‘huge’	5	0	2.5
	2 – ‘big’	0	0	0
	1 – ‘small’	95	100	97.5

Katsos & Bishop, 2011

Sentence-to-Picture matching

- 5-to-6-year-old children
- Stimuli are similar to experiment 1 and 2 but participants are asked to select the picture that best matches the critical sentence
- If children are not competent with informativeness, they should pick the picture that matches the informative or the under-informative interpretation at chance

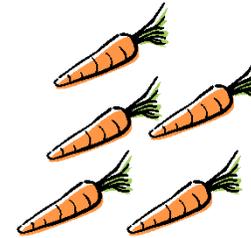




This one!
87%



4%



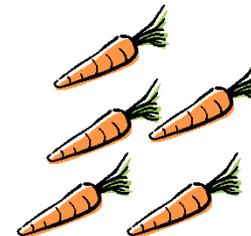
The mouse picked up some of the carrots



6%



3%



- Sentence evaluation tasks systematically misrepresent children's pragmatic competence

- Are children tolerant of violations in general?
No, they always reject semantically false utterances. Tolerance applies to violations of pragmatics only.
- Why should *children* be tolerant to pragmatic violations?
Some adults are tolerant too! (recall, 28% of adult responses were hedges rather than downright rejections)
- But why is this tolerance so pronounced in children?
Children may be less confident about their meta-linguistically judgments

- Adults are not that different



- Qualitative evidence:

Revisions and hedges:

“Yes, ... all the carrots / Yes, ...but the heart as well”

“Not wrong...not right either...” / ‘Half right!’ / “I can’t say”

- **This study: reaction times**

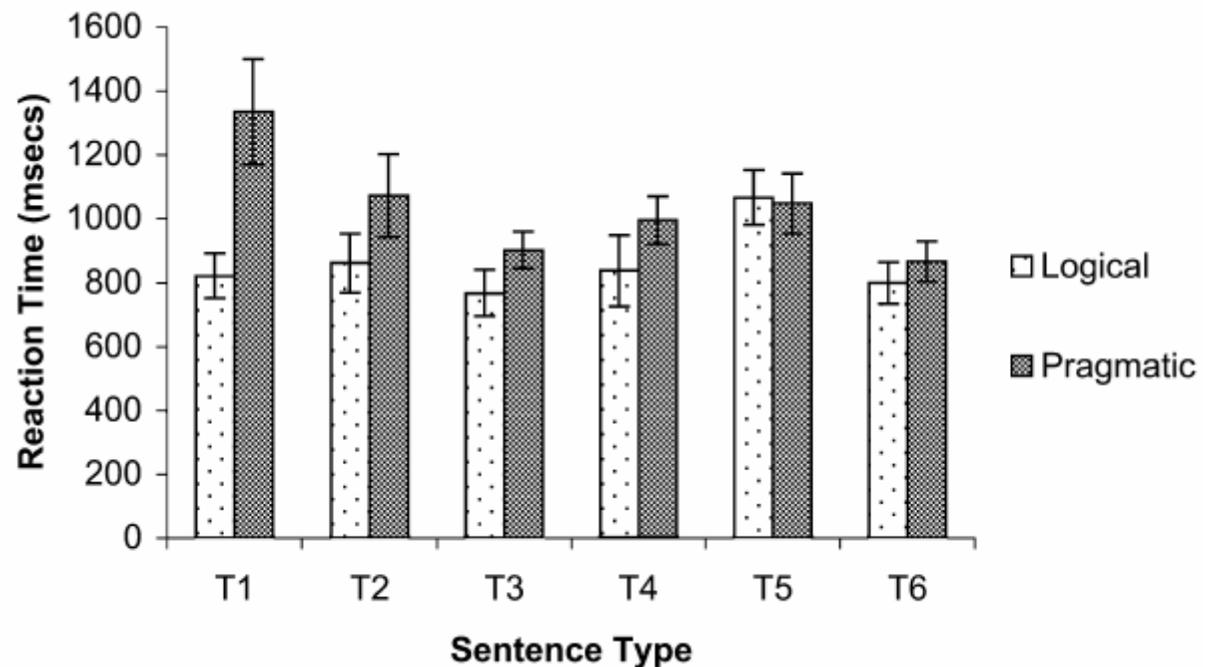
– Acceptance of true and infelicitous utterances should be slower than acceptance of true and felicitous utterances

Prior work?

- Bott & Noveck 2004, Tomlinson et al. 2013, Bott et al. 2012 no such effect
- E.g. in Bott & Noveck 2004 Bott & Noveck 2004, T1 = T2 for Logical responders

Table 1
Examples of the sentence types used in Experiments 1–4

Reference	Example sentence
T1	Some elephants are mammals
T2	Some mammals are elephants
T3	Some elephants are insects
T4	All elephants are mammals
T5	All mammals are elephants
T6	All elephants are insects

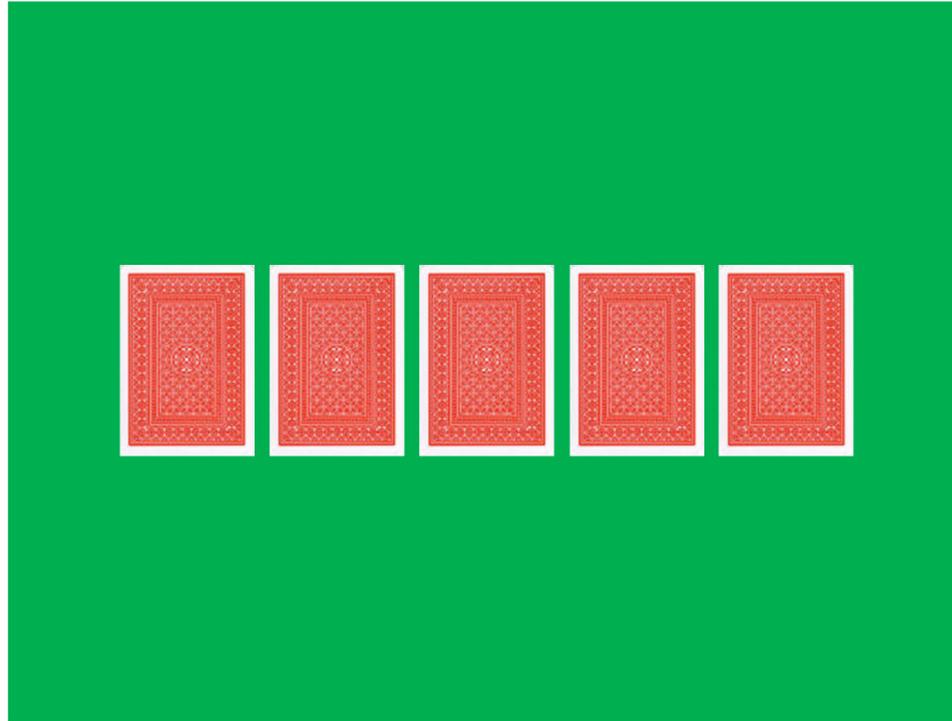


- But participants were instructed to interpret 'some' as 'at least some' and then given training



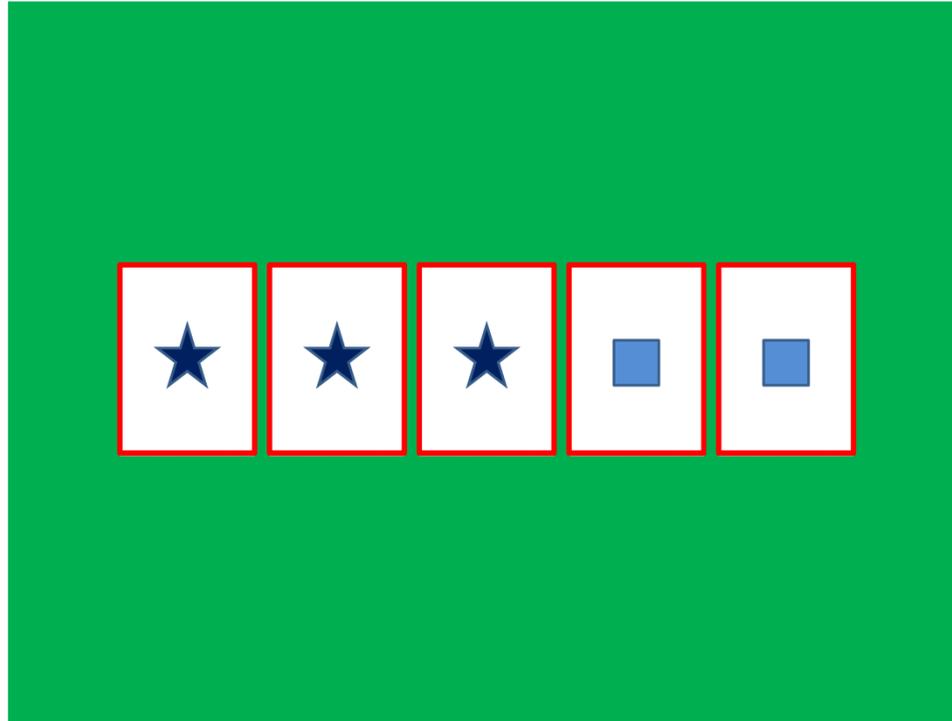
Antoniou, Cummins & Katsos (in prep)

Sentence evaluation task

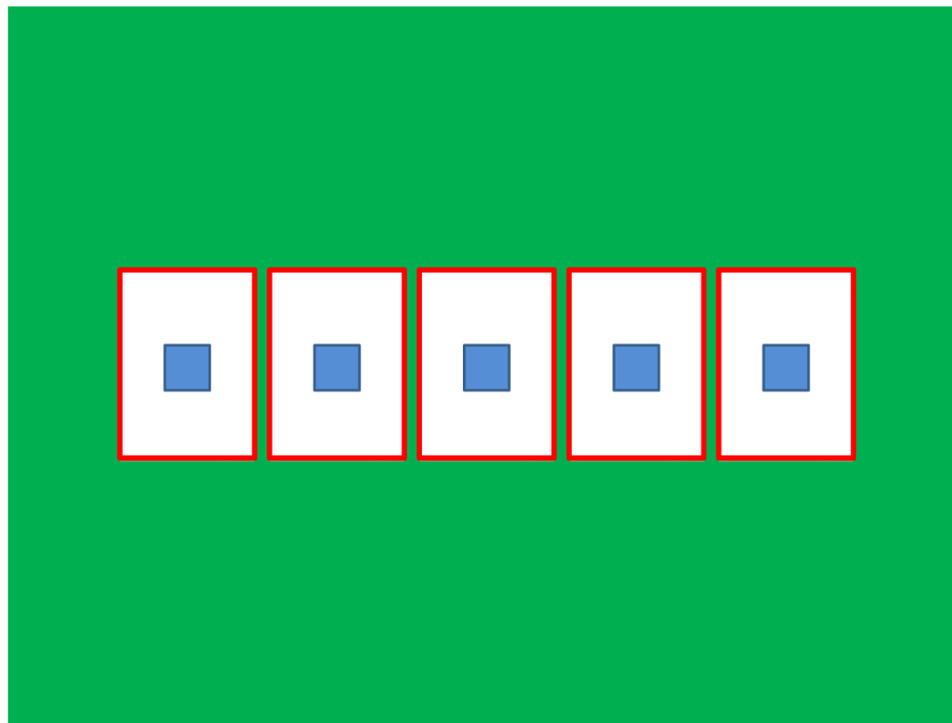


Antoniou, Cummins & Katsos (in prep)

Some of the cards have a star



Some of the cards have a star



40 participants, native speakers of English (aged 18-40, mean 23.1, SD 4.5 years; 27 female)

Figure 5: Number of participants per rate of rejection of under-informative items in experiment 2.

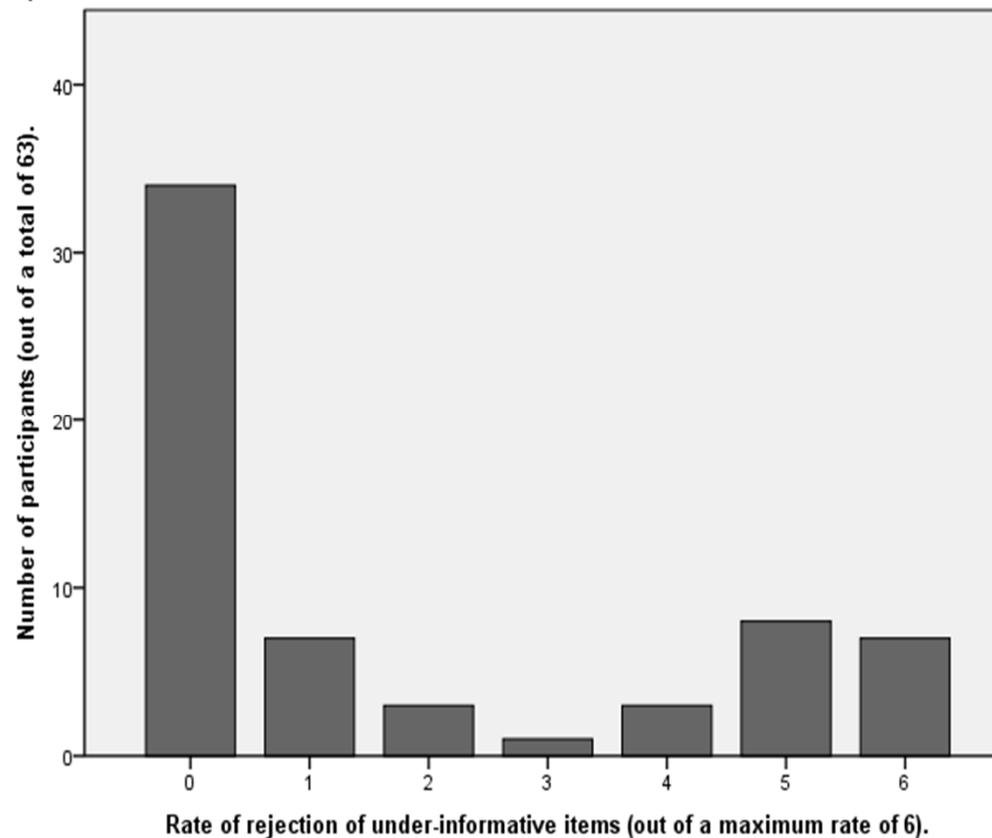
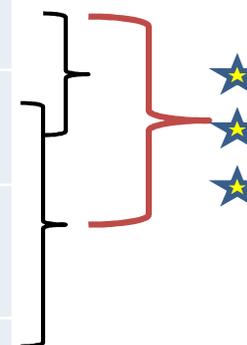


Table 2: Test conditions and summary data (excludes erroneous and slow (>5s) responses) for experiment 2

Quantifier	Condition	Condition status	Expected response(s)	Rate of response (%)	Reaction time, SD (ms)
Some	5/5	True but under-informative	True	68.0%	1191, 864
			False	28.8%	1429, 893
	3/5	True and informative	True	98.1%	905, 401
	0/5	False	False	99.2%	1108, 553
All	5/5	True and informative	True	98.4%	879, 488
	3/5	False	False	97.6%	1000, 441
None	3/5	False	False	95.9%	1337, 925
	0/5	True and informative	True	89.7%	1298, 681



RTs of true responses for under-informative some (313 responses, mean RT 1158ms, SD 731ms) compared to those for fully informative some (447 responses, mean RT 822ms, SD 444ms).

coefficient =452.86, SE=90.60, t=4.998, p<.001

Are these effects general to all participants?

- Comparing the 3/5 and 5/5 cases, 37 of the 39 participants were faster in the 3/5 case ($p < 0.001$, sign test).
- Of the 16 who consistently accepted *some* in the 5/5 case (accepting it five or six out of six trials), 14 responded slower than in the 3/5 case (sign test, $p < 0.01$).

Interim Summary

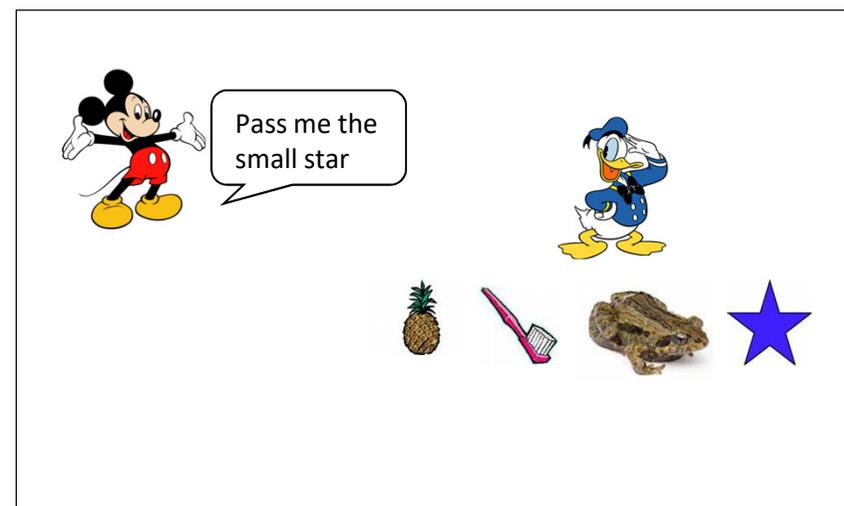
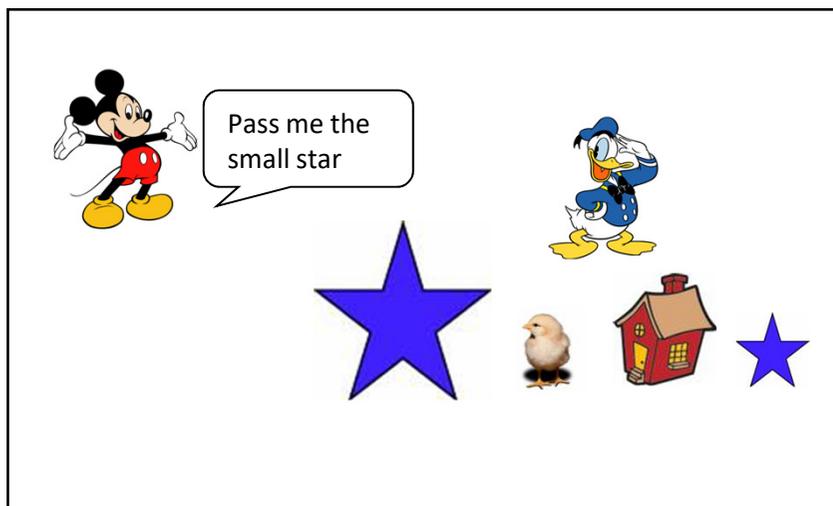
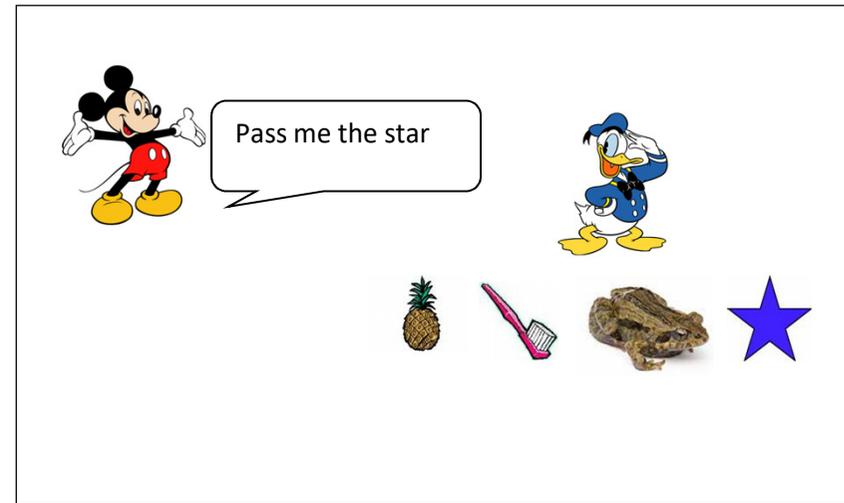
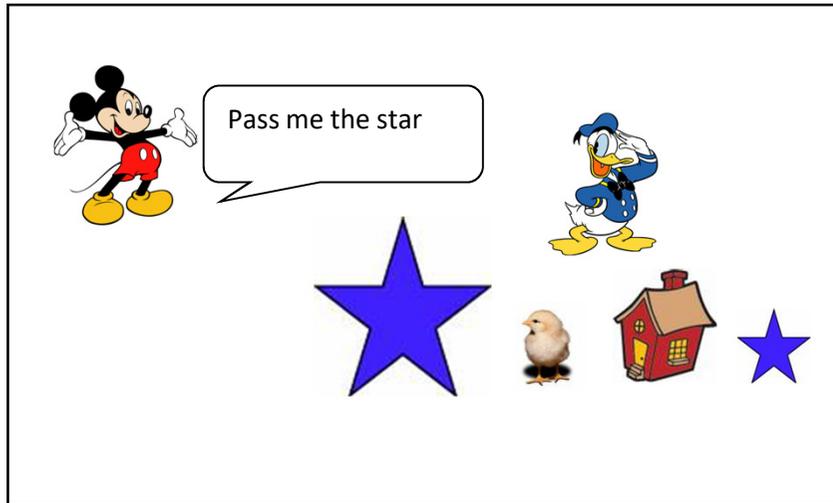
Children from 5 ½ years-old onwards are tolerant little Grices: they do not consider violations of informativeness to be grave enough to warrant downright rejection of the critical utterance

Adults are tolerant too (qualitative and quantitative data)

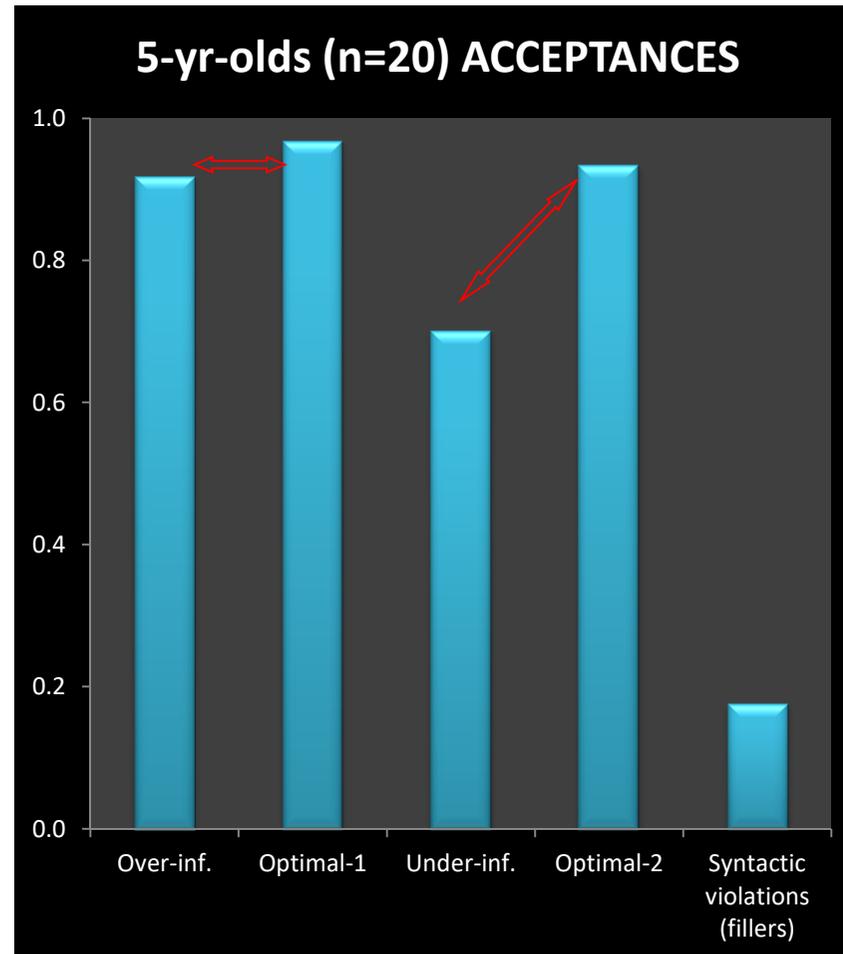


Davies & Katsos, 2010: over- and under-informativeness in referring expressions

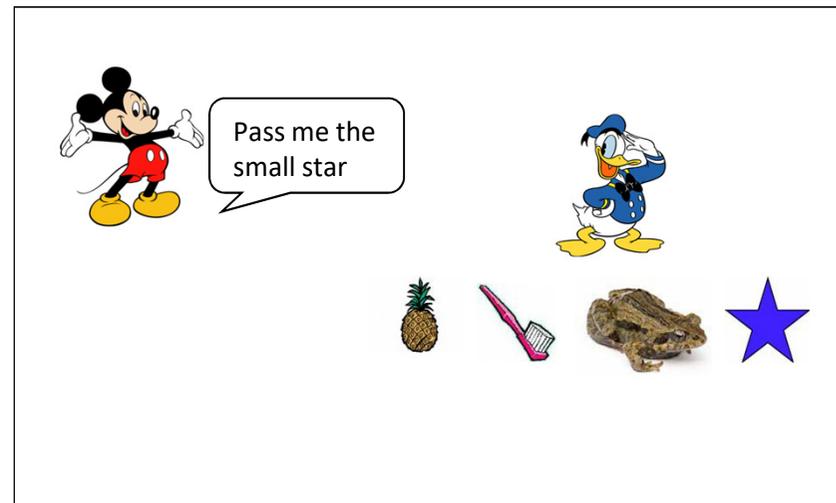
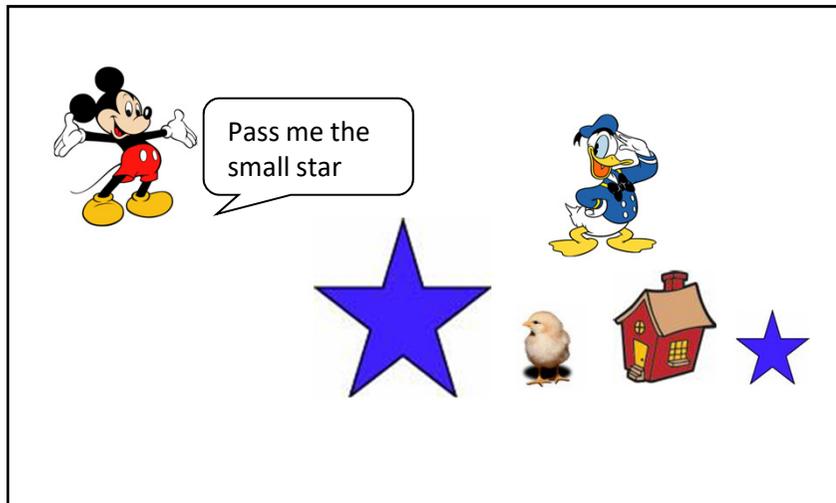
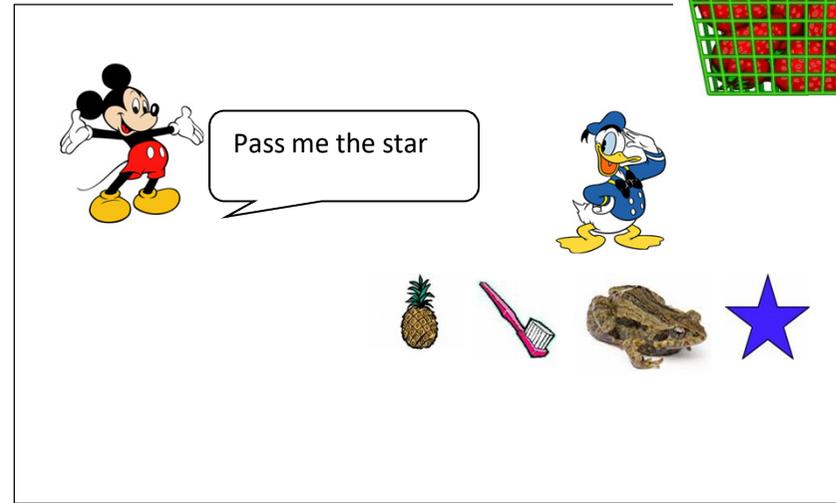
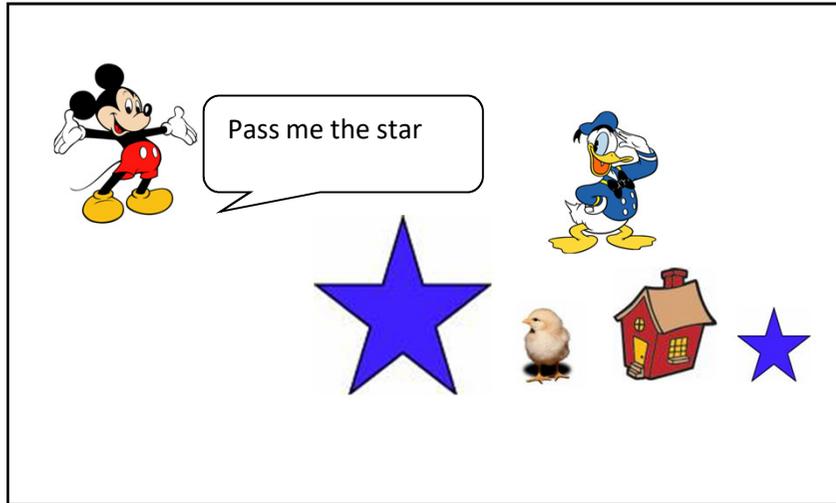
Davies & Katsos, 2010: over- and under-informativeness in referring expressions



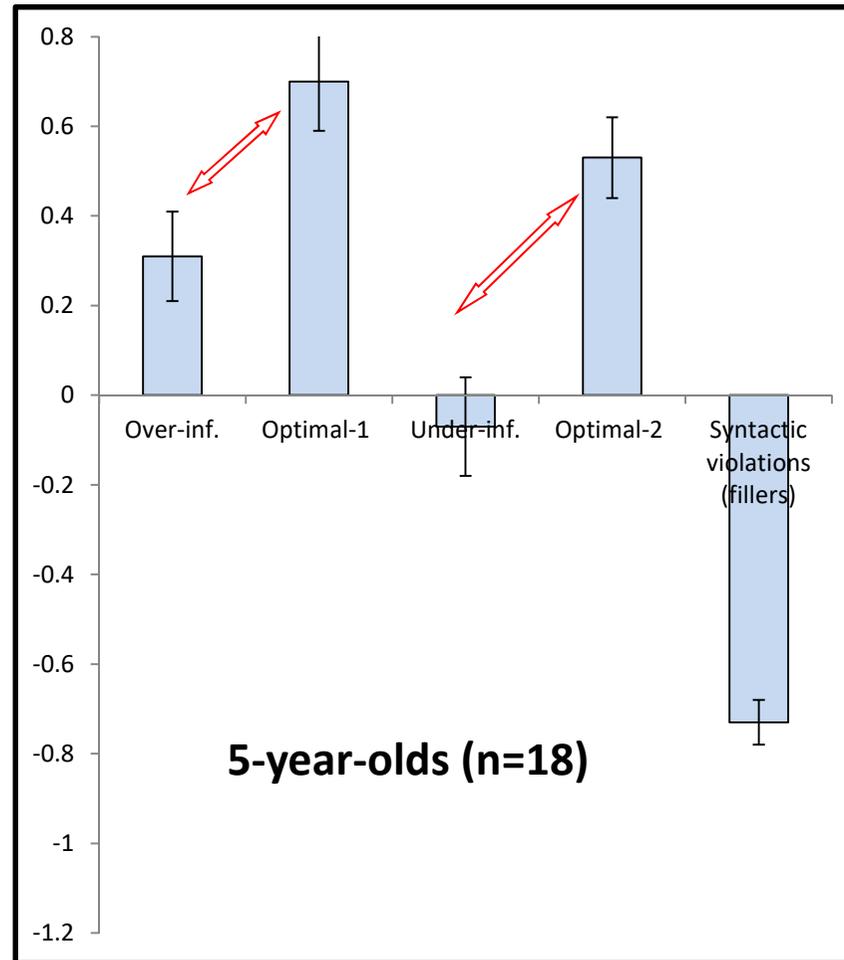
Expt 2a: Binary judgments: mean acceptances by 5-year-olds



Comprehension studies: experimental conditions



Expt 2b: Magnitude Estimation (z-transformed): mean ratings by 5-year-olds.



All within-group comparisons are significantly different ($p < .001$ - $p < .05$), unless otherwise marked.

Interim Summary

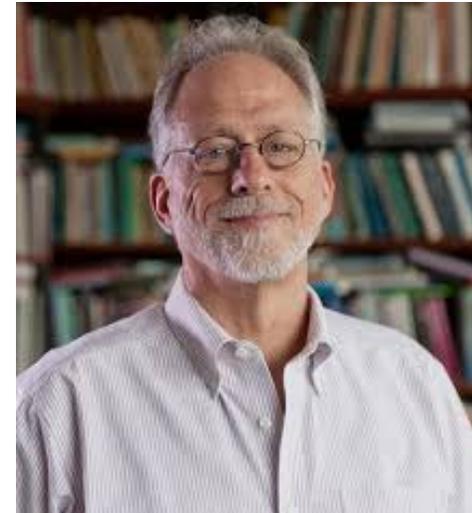
In sentence judgement tasks:

Children from 5 ½ years-old onwards are tolerant little Grices on a number of maxims

Adults are tolerant too

But where have we got to with regards to the theoretical question?

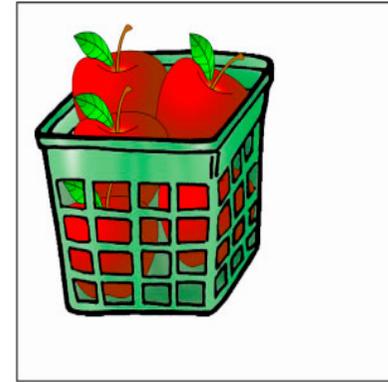
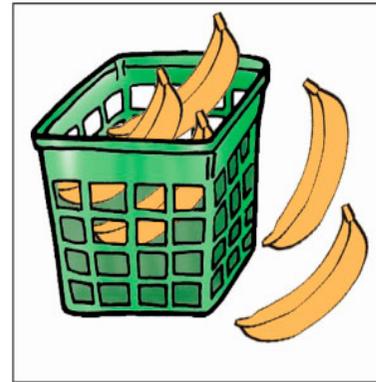
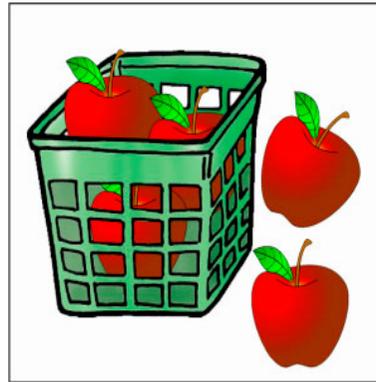
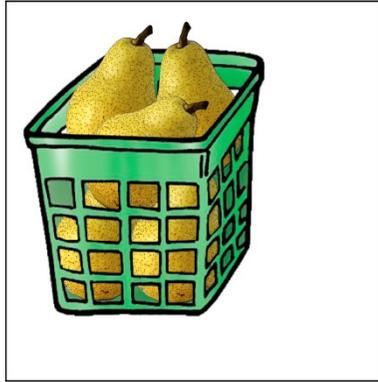
Nowhere really, because children performed as well with the scalars as with the ad hocs in experiment 2 and 3 of Katsos & Bishop, 2011, and in both cases at ceiling rates.



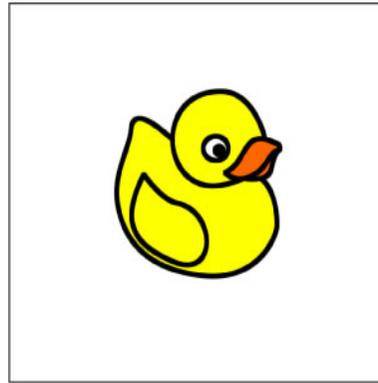
Grosse, Schulze, Tomasello & Katsos (submitted)

- German-speaking participants:
 - 3-1/2-year-olds (n=24; 12 female)
 - 5-year-olds (n=24; 12 female)
 - Adults (n=12; 6 female)

Another attempt: hiding game



Grosse, Schulze, Tomasello & Katsos (submitted)



Grosse, Schulze, Tomasello & Katsos (submitted)

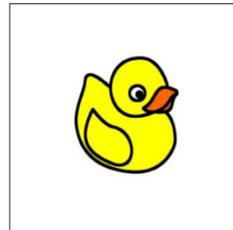
Doll and car



Car



Duck



Bucket



Ad hoc

	Contrast	Test	Control	Under-info
3y	0.93 ^{***}	0.81 ^{***}	0.96	0.97
5y	0.97 ^{***}	0.97 ^{***}	1.00	1.00
adults	1.00	0.86 ^{***}	1.00	1.00

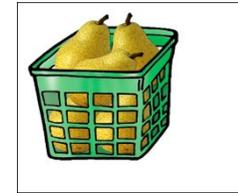
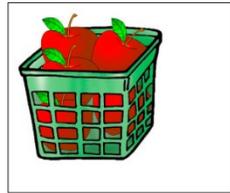
Grosse, Schulze, Tomasello & Katsos (submitted)

All apples

Some apples

Some bananas

Some pears



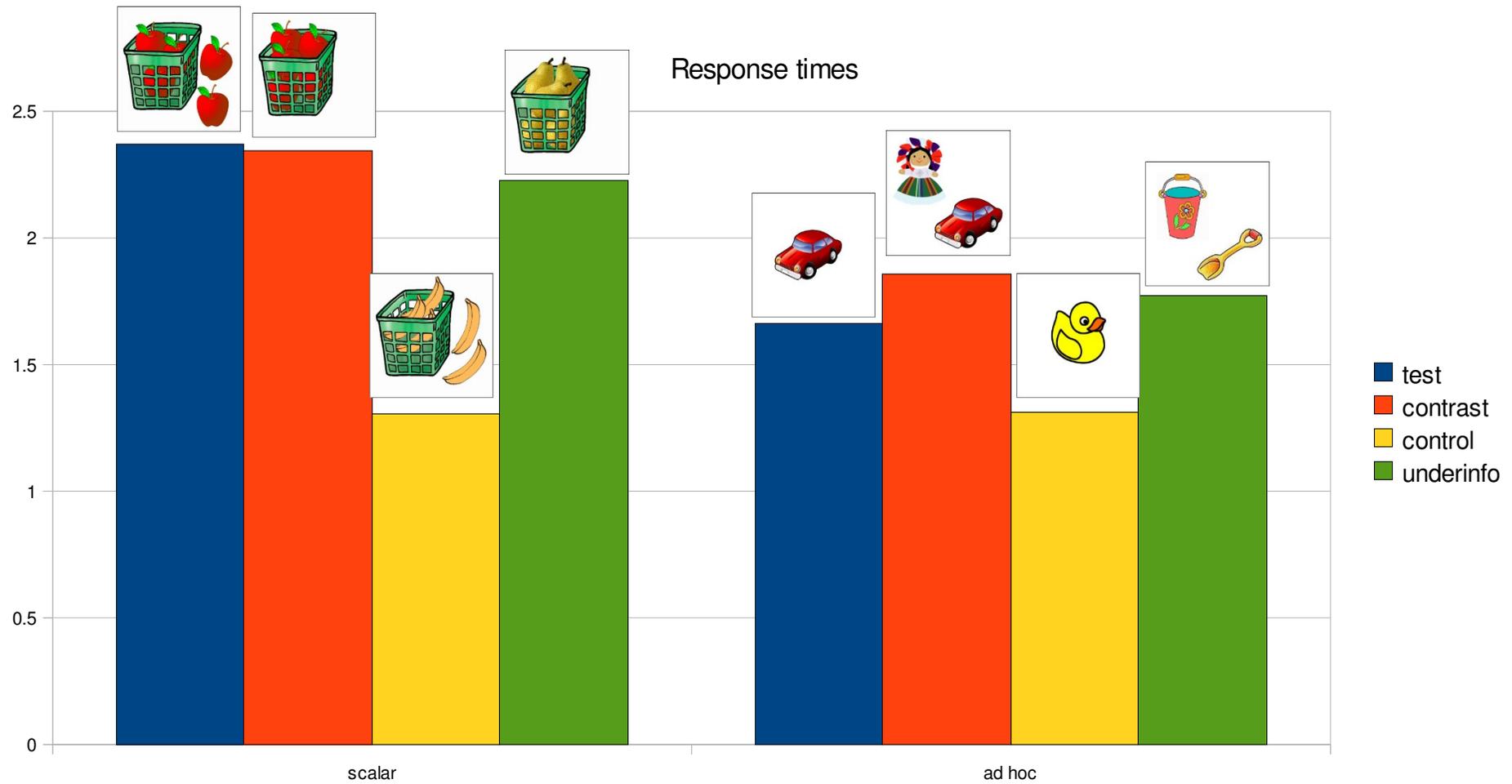
Scalar

	Contrast	Test	Control	Under-info
3y	0.61	0.48	0.99	0.92
5y	0.92 ^{***}	0.69 ^{**}	0.99	0.97
adults	1.00 ^{***}	0.86 ^{***}	1.00	1.00

Grosse, Schulze, Tomasello & Katsos (submitted)

2 (Felicity: Control v Under-info) x 2 (Type: Scalar v Adhoc) ANOVA

Felicity $F(1, 23) = 22,48, ***$; Type $F(1, 23) = 1,72, n.s.$ Felicity x Type $F(1, 23) = 1,31, n.s.$



Reaction times: 3- and 5-year olds seem sensitive to under-informativeness

Box selection:

3-year-olds: succeed with ad hoc but not scalar

5-year-olds: succeed with both

Correlational analyses for box selection:

- 'some' correlates with 'all' ($r = .41 *$)
- 'NP' correlates with 'NP & NP' ($r = .41 *$)

Performance with the stronger contrast-mate correlates with performance with the critical weaker contrast-mate

- Age at which children 'do implicatures' much younger than suspected.

Stiller, Goodman & Frank, 2014

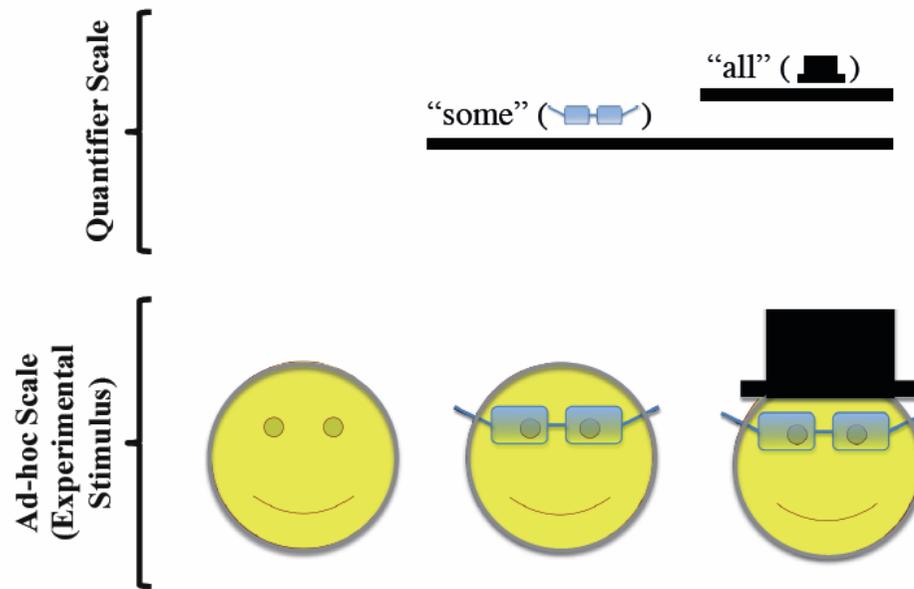


Figure 1. Example stimuli from our referent selection task. The middle item represents the pragmatically enriched interpretation of “My friend has glasses.” The implicature has a similar logical structure to the conventional *some-not-all* implicature (top).

		Experimental (Label)			Control (No-Label)			
Age	n	Mean	Std. Err.	Std. Dev	n	Mean	Std. Err.	Std. Dev
2 years	25	0.57	0.05	0.50	24	0.24	0.04	0.43
3 years	26	0.71	0.04	0.46	24	0.24	0.04	0.43
4 years	24	0.77	0.04	0.42	24	0.13	0.03	0.33
Adults	24	0.96	0.02	0.20	24	0.23	0.04	0.42

Conclusion

- Several theories of scalar implicature
- Several similarities between scalar implicature and other key competencies (word learning by exclusion)
- Developmentally, we have progressed from kids failing implicatures at the age of 8 to kids passing some implicatures at the age of 3 ½, very close to bridging the gap between implicature and word-learning
- Sentence evaluation tasks systematically misrepresent child and adult pragmatic competence

Thank you!

Tomorrow's lecture:

Quantity implicature and its relatives

Napoleon Katsos

nk248@cam.ac.uk