

Learnability of Quantifiers

Developmental psychology and unification



Recap

Recap

- Yesterday:
 - Neural network notebook
 - Ease of learning explains universals for quantifiers, responsive predicates, and color terms

Recap

- Yesterday:
 - Neural network notebook
 - Ease of learning explains universals for quantifiers, responsive predicates, and color terms
- Today:
 - Developmental psychology of quantifiers
 - Cross-linguistic picture
 - Discussion

Conservativity and Learnability of Determiners

TIM HUNTER
Yale University

JEFFREY LIDZ
University of Maryland

Abstract

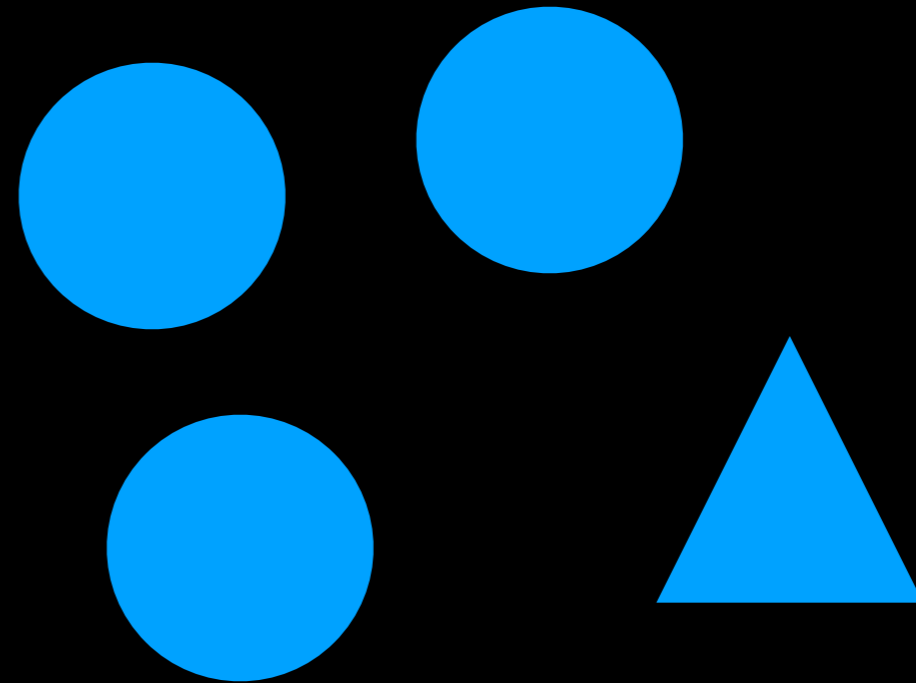
A striking cross-linguistic generalisation about the semantics of determiners is that they never express non-conservative relations. To account for this one might hypothesise that the mechanisms underlying human language acquisition are unsuited to non-conservative determiner meanings. We present experimental evidence that 4- and 5-year-olds fail to learn a novel non-conservative determiner but succeed in learning a comparable conservative determiner, consistent with the learnability hypothesis.

Children's symmetric interpretation of quantifiers

Children's symmetric interpretation of quantifiers

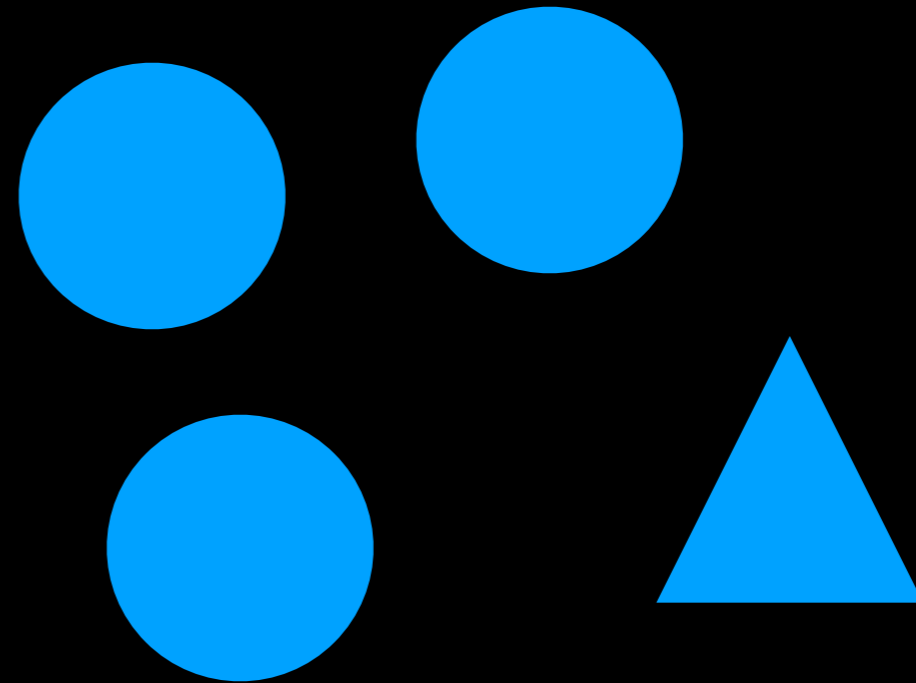
1. Are all the circles blue?

Children's symmetric interpretation of quantifiers



1. Are all the circles blue?

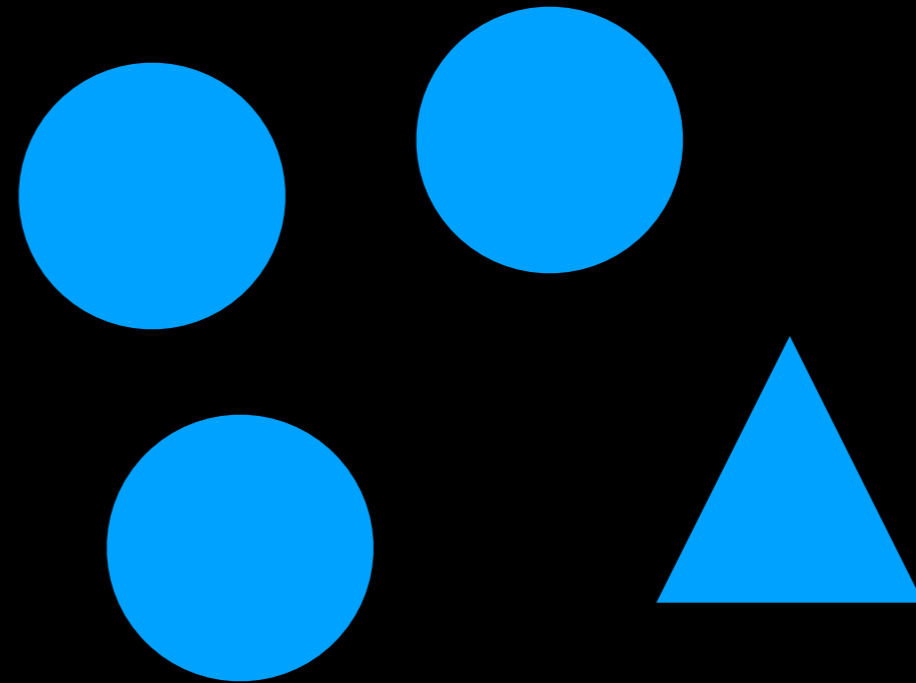
Children's symmetric interpretation of quantifiers



1. Are all the circles blue?

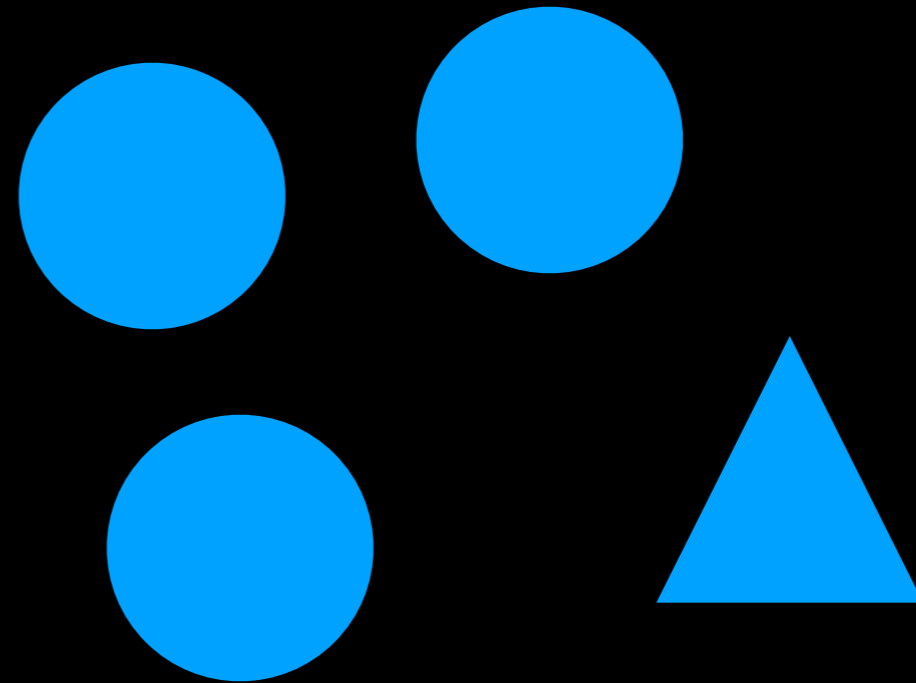
- All(A, B) iff $A = B$

Children's symmetric interpretation of quantifiers



1. Are all the circles blue?
 - $All(A, B) \text{ iff } A = B$
2. Is every girl riding an elephant?

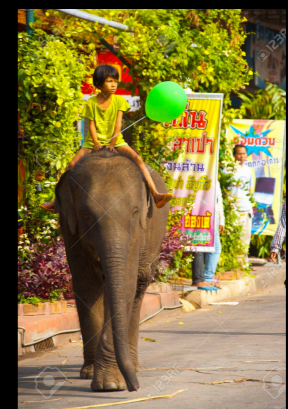
Children's symmetric interpretation of quantifiers



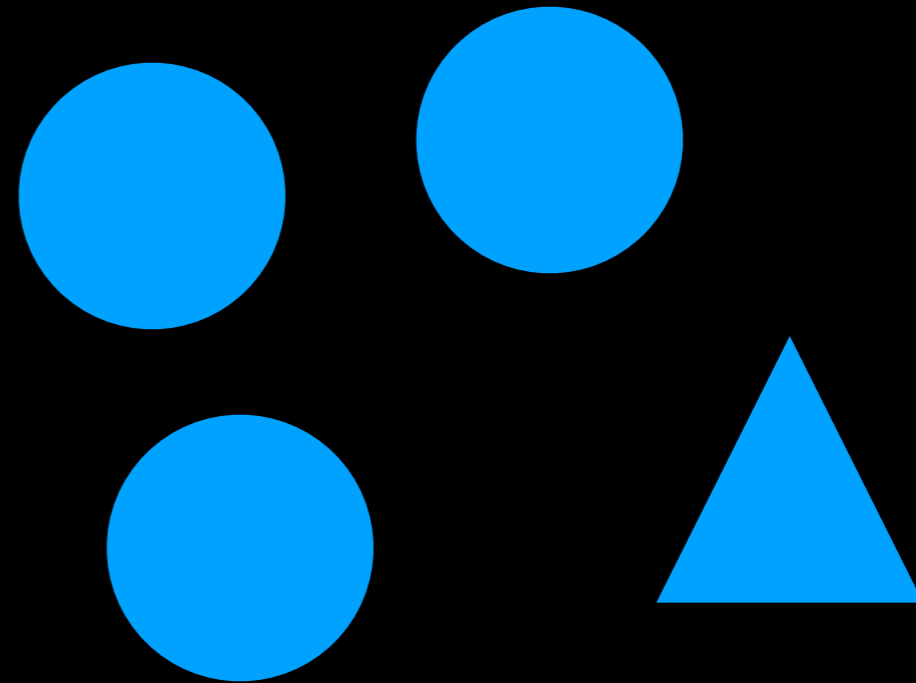
1. Are all the circles blue?

- $\text{All}(A, B) \text{ iff } A = B$

2. Is every girl riding an elephant?



Children's symmetric interpretation of quantifiers



1. Are all the circles blue?

- $\text{All}(A, B) \text{ iff } A = B$

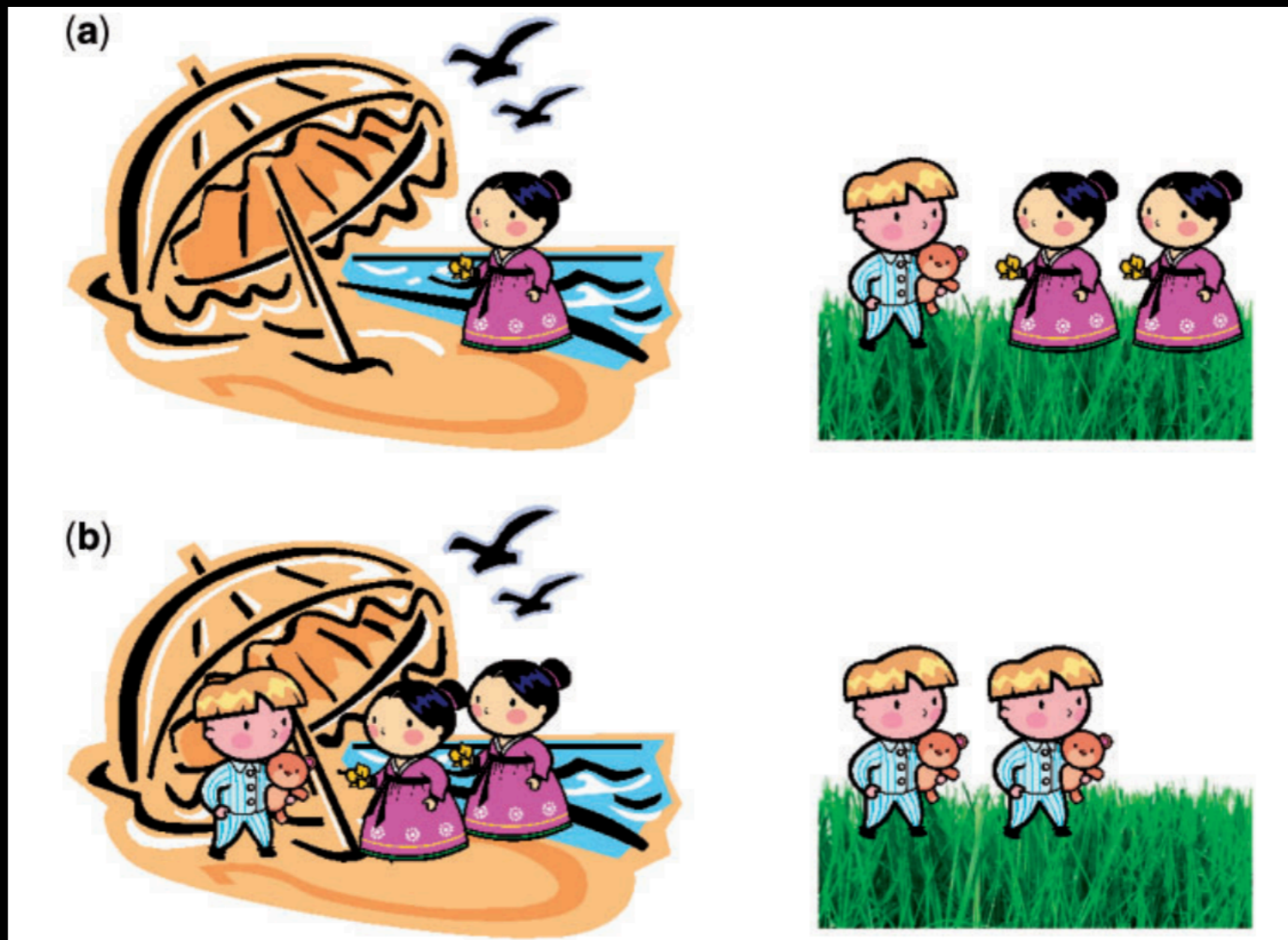
2. Is every girl riding an elephant?

- Elephant-riders not elephants



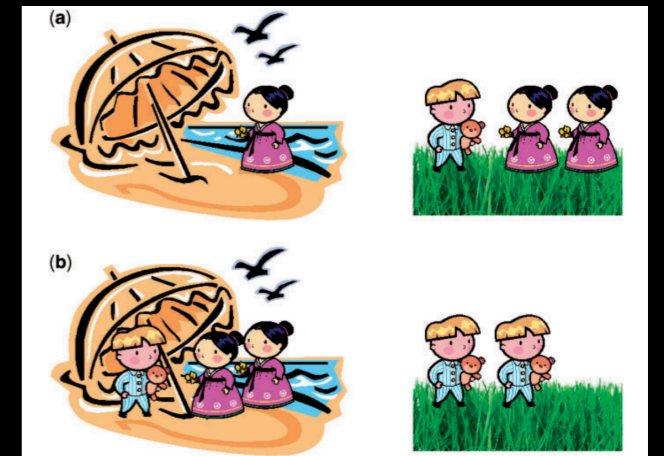
Gleeb vs Gleeb'

- Gleeb = 'not all'
- Gleeb' = 'not only'



Lidz & Hunter's experiment

- Two conditions: CONS and non-CONS
- Picky puppet task (Waxman & Gelman 1986).
- Warm-up (3 cards) - Training (5 cards) - Target (5 cards)
 1. *The puppet told me that he likes this card because gleeb girls are on the beach*
 2. *The puppet told me that he doesn't like this card because it not true that gleeb girls are on the beach.*

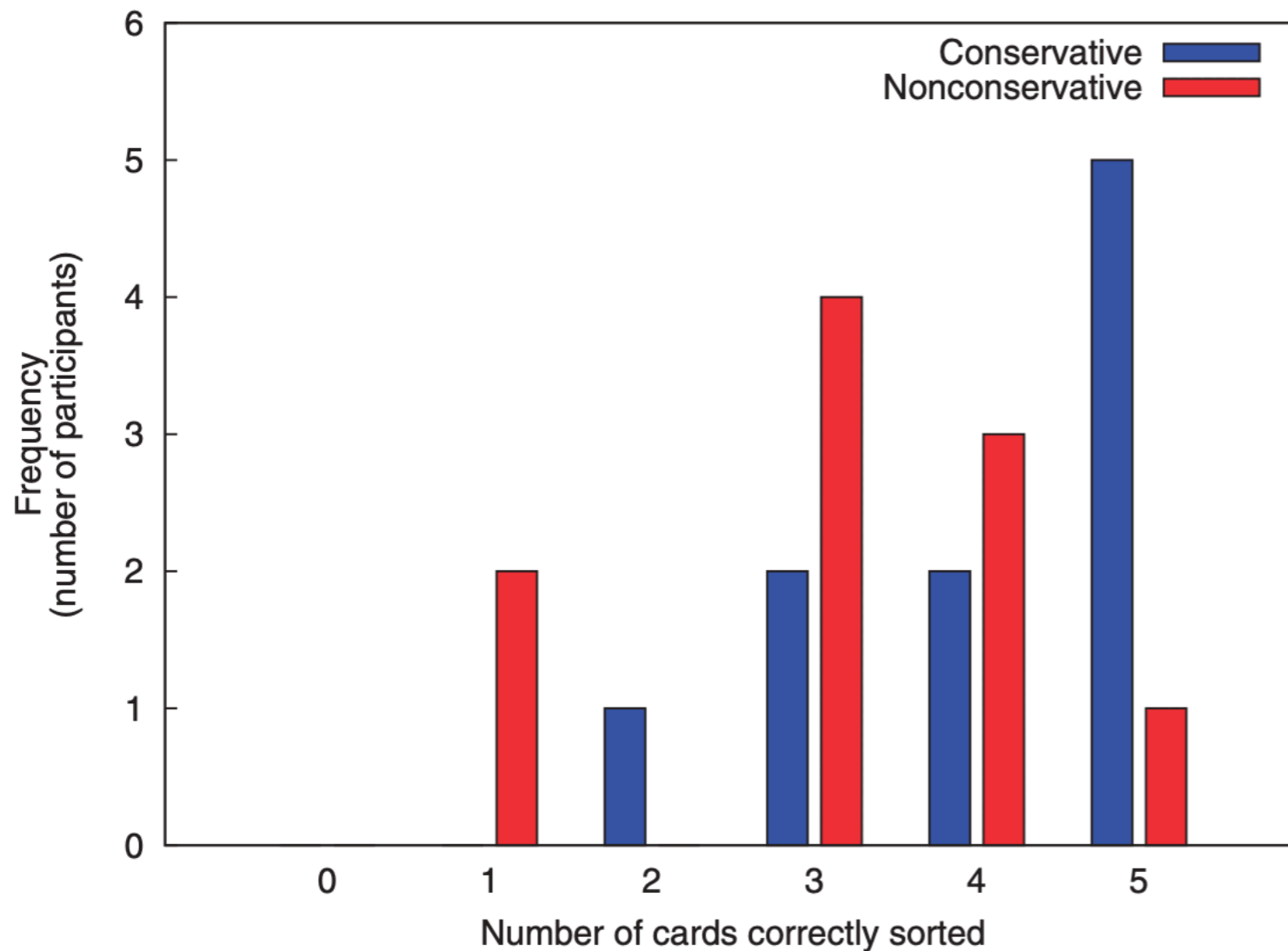


Participants

- 20 children
- Aged 4.5 to 5.6 (mean 5.0)
- Conservative condition 4.5 to 5.5 (mean 4.11)
- Non-conservative condition 4.11 to 5.3 (mean 5.1)

Results

Condition	Conservative	Non-conservative
Cards correctly sorted (out of 5)	mean 4.1 (above chance, $p < 0.0001$)	mean 3.1 (not above chance, $p > 0.2488$)
Subjects with “perfect” accuracy	50%	10%



**“the puppet was confused about which characters
on the cards were boys and which were girls”**

-Perfect non-conservative kid

Discussion

Discussion

- Are kids interpreting *gleeb* as a GQ?

Discussion

- Are kids interpreting *gleeb* as a GQ?
- Are the results consistent with a structural account?

Discussion

- Are kids interpreting *gleeb* as a GQ?
- Are the results consistent with a structural account?
- Do the *gleeb* and *gleeb'* sentences differ only in conservativity?

Discussion

- Are kids interpreting *gleeb* as a GQ?
- Are the results consistent with a structural account?
- Do the *gleeb* and *gleeb'* sentences differ only in conservativity?
- Would that replicate? At a larger scale?

Discussion

- Are kids interpreting *gleeb* as a GQ?
- Are the results consistent with a structural account?
- Do the *gleeb* and *gleeb'* sentences differ only in conservativity?
- Would that replicate? At a larger scale?
- Also, what about the other universals?

Connecting Content and Logical Words

Emmanuel Chemla and Brian Buccola

Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, ENS, EHESS, CNRS, PSL University, Paris, France

Email: chemla@ens.fr, brian.buccola@gmail.com

Isabelle Dautriche

Centre for Language Evolution, University of Edinburgh, Edinburgh, UK

Email: isabelle.dautriche@gmail.com

First version received 9 February 2017; Second version received 19 August 2018; Accepted 21 August 2018

Abstract

Content words (e.g. nouns and adjectives) are generally connected: there are no gaps in their denotations; no noun means 'table or shoe' or 'animal or house'. We explore a formulation of connectedness which is applicable to content and logical words alike, and which compares well with the classic notion of monotonicity for quantifiers. On

“No noun in English means ‘bottle or eagle’, and no quantifier means ‘less than 5 or more than 10’.”

Connectedness

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1] : ta + (1 - t)b \in X$

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1] : ta + (1 - t)b \in X$
- c is between a and b : $\exists t \in [0,1] : c = ta + (1 - t)b$

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1] : ta + (1 - t)b \in X$
- c is between a and b : $\exists t \in [0,1] : c = ta + (1 - t)b$
- In general: betweenness can be 'primitive'

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1] : ta + (1 - t)b \in X$
- c is between a and b : $\exists t \in [0,1] : c = ta + (1 - t)b$
- In general: betweenness can be 'primitive'
- f is **connected**: if c is between a and b , then $f(c) \geq f(a)$ or $f(c) \geq f(b)$

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1] : ta + (1 - t)b \in X$
- c is between a and b : $\exists t \in [0,1] : c = ta + (1 - t)b$
- In general: betweenness can be ‘primitive’
- f is **connected**: if c is between a and b , then $f(c) \geq f(a)$ or $f(c) \geq f(b)$
- Exercise: X is convex iff its characteristic function is connected

Connected Quantifiers

Connected Quantifiers

- C is between A and B : $A \subset C \subset B$ (or vice versa)

Connected Quantifiers

- C is between A and B : $A \subset C \subset B$ (or vice versa)
- So: Q (of type $\langle et, t \rangle$) is connected:
if $A \subset C \subset B$ then $Q(C) \geq Q(A)$ or $Q(C) \geq Q(A)$

Connected Quantifiers

- C is between A and B : $A \subset C \subset B$ (or vice versa)
- So: Q (of type $\langle et, t \rangle$) is connected:
if $A \subset C \subset B$ then $Q(C) \geq Q(A)$ or $Q(C) \geq Q(B)$
- Exercise: show that the above is equivalent to:
if $A \subset C \subset B$ and $Q(A) = 1$ and $Q(B) = 1$, then
 $Q(C) = 1$

Connected Quantifiers

- C is between A and B : $A \subset C \subset B$ (or vice versa)
- So: Q (of type $\langle et, t \rangle$) is connected:
if $A \subset C \subset B$ then $Q(C) \geq Q(A)$ or $Q(C) \geq Q(B)$
- Exercise: show that the above is equivalent to:
if $A \subset C \subset B$ and $Q(A) = 1$ and $Q(B) = 1$, then
 $Q(C) = 1$
- Note: this property has also been called *continuity* by van Benthem (1984, 1986)

Theorem. Q is monotone iff Q and $\neg Q$ are connected.

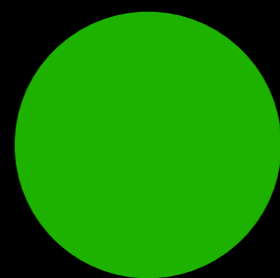
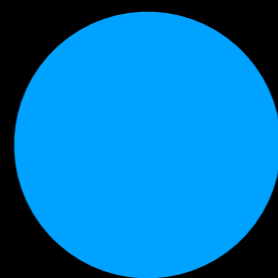
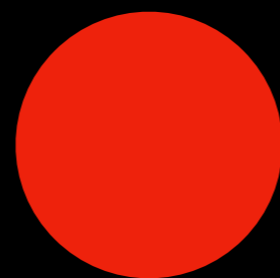
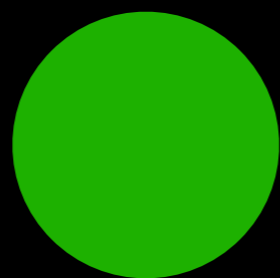
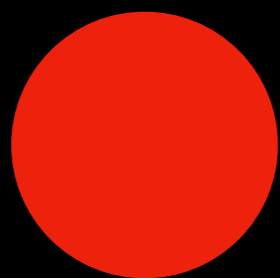
All word meanings are connected. (?)

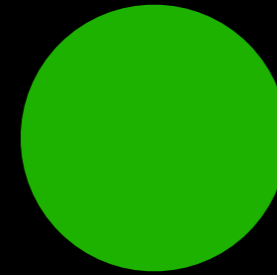
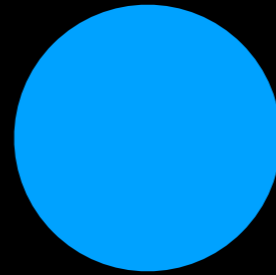
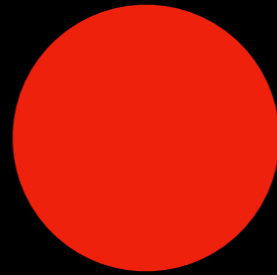
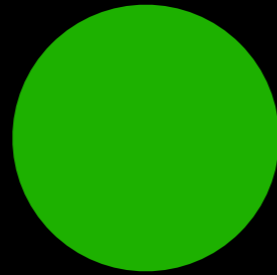
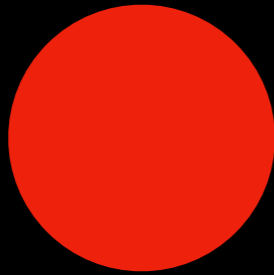
“Connected quantifiers are typically expressed in simpler ways than non-connected quantifiers.”

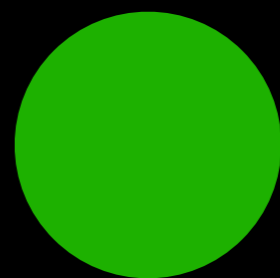
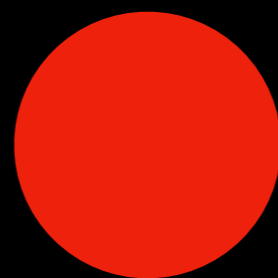
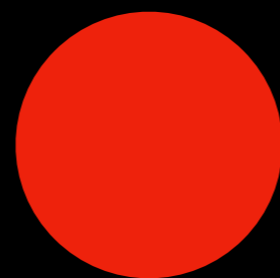
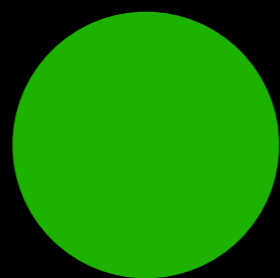
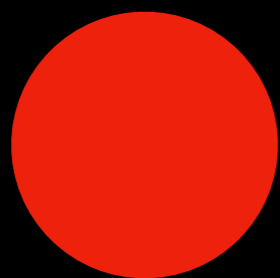
“Connected quantifiers are typically expressed in simpler ways than non-connected quantifiers.”

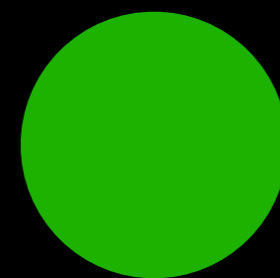
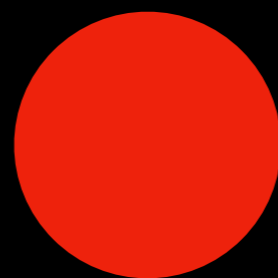
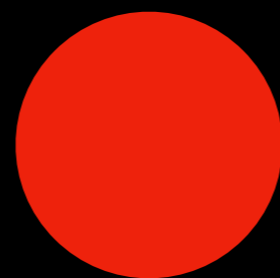
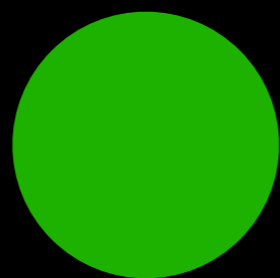
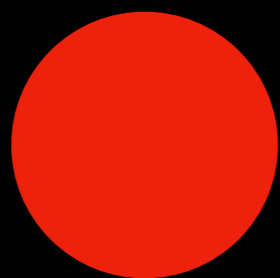
Compare, e.g., ‘between 5 and 10’ or ‘5 to 10’ with ‘less than 5 or more than 10’.

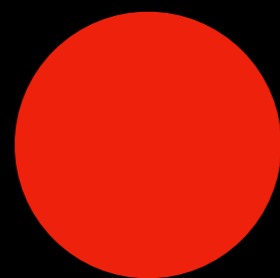
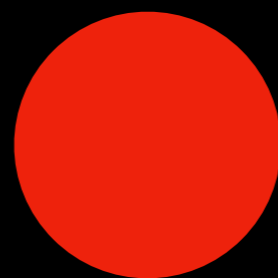
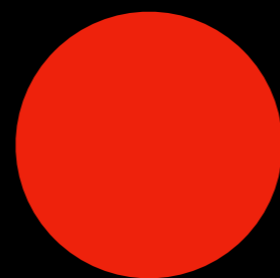
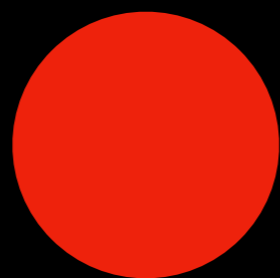
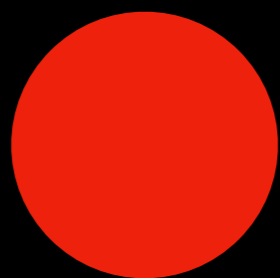
Experiment Time!

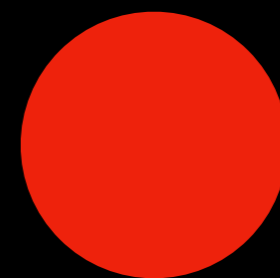
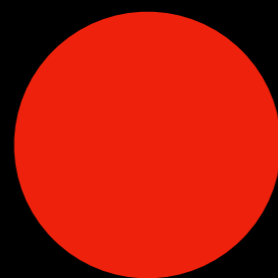
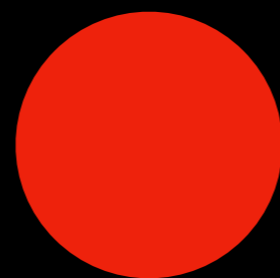
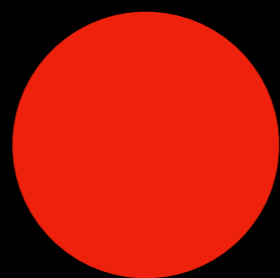
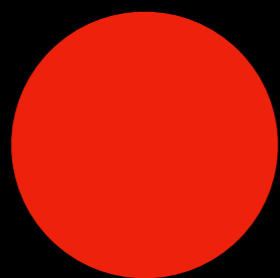


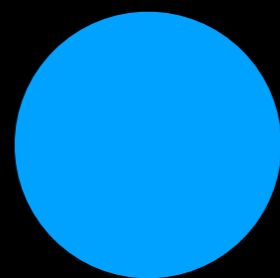
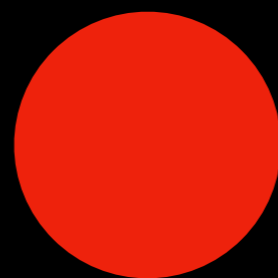
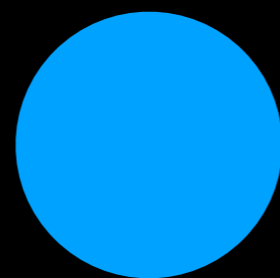
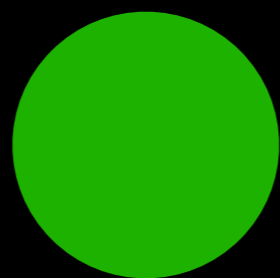
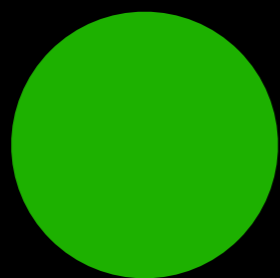


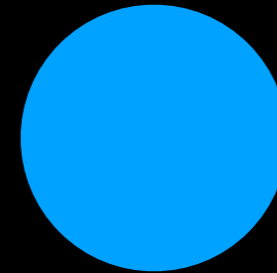
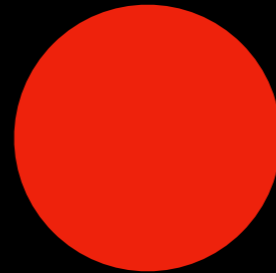
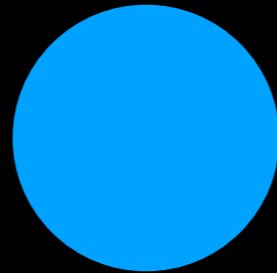
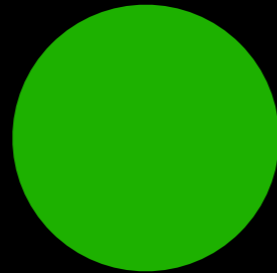
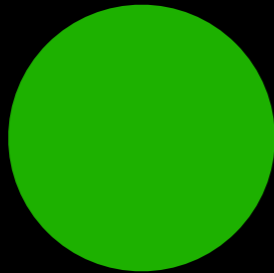


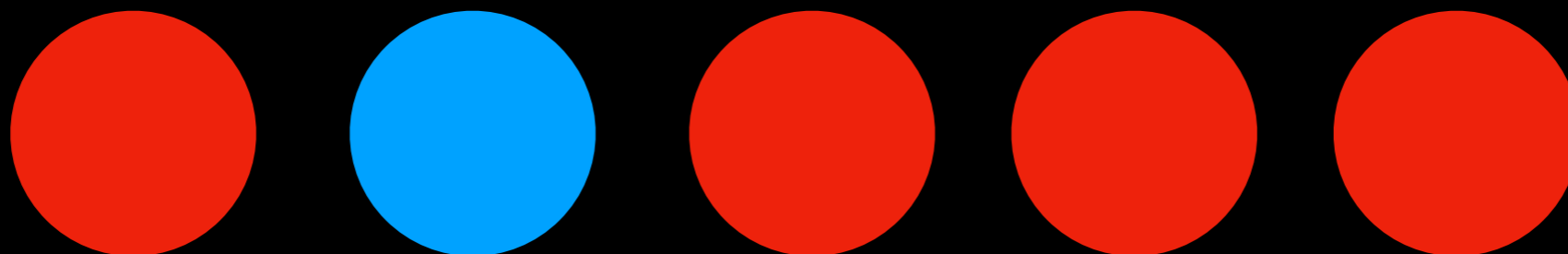


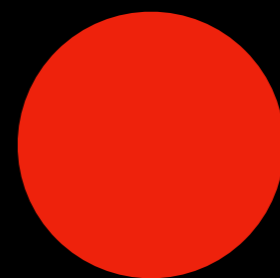
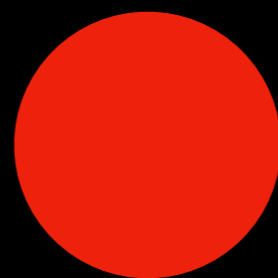
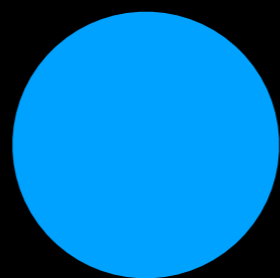
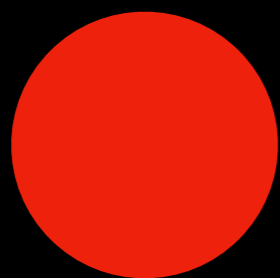


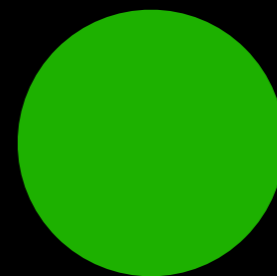
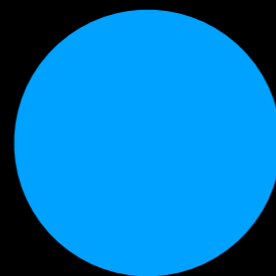
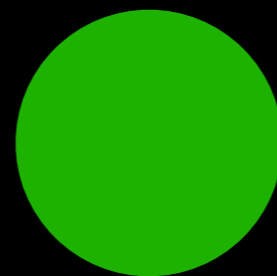
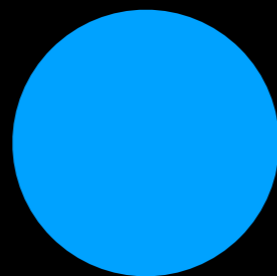
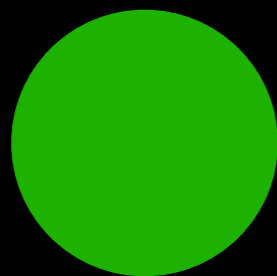


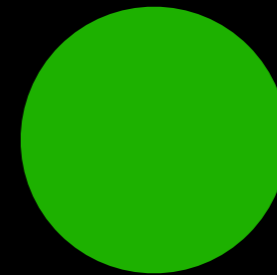
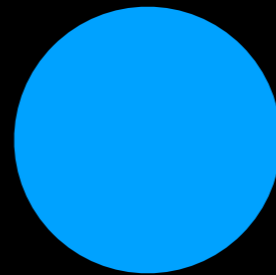
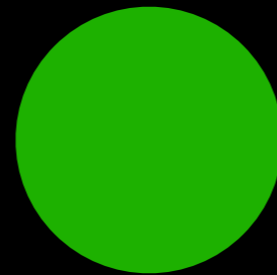
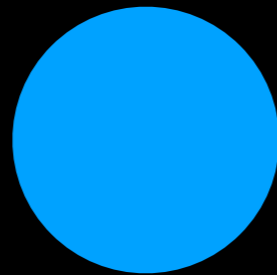
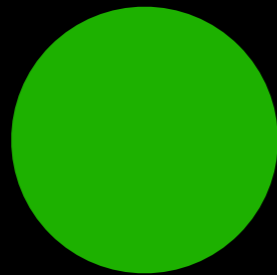






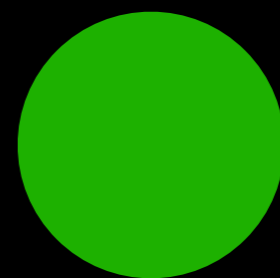
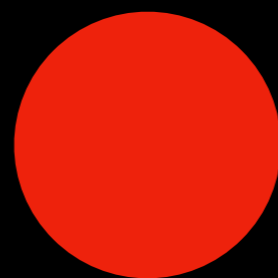
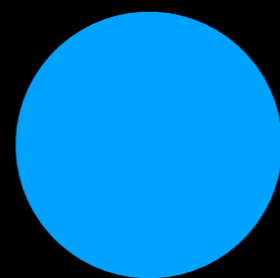
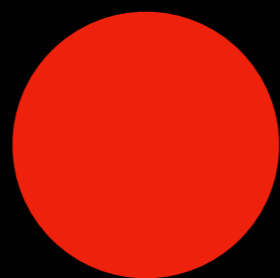


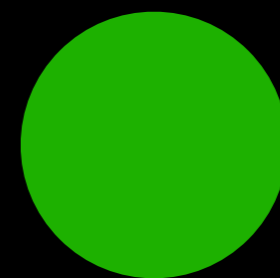
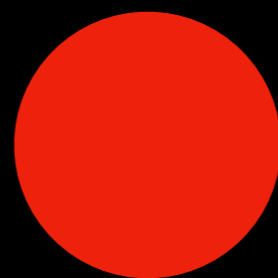
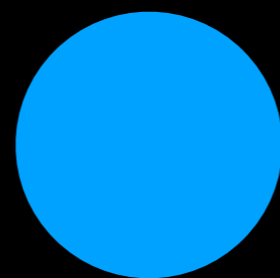
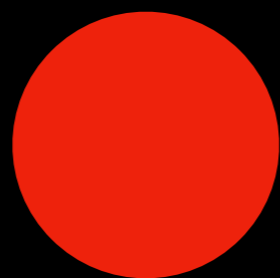
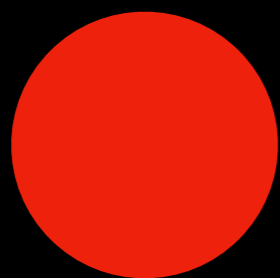


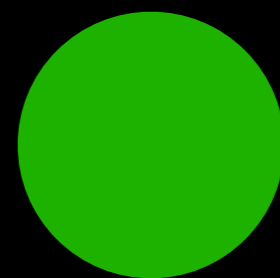
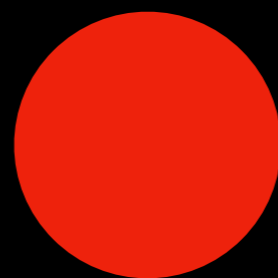
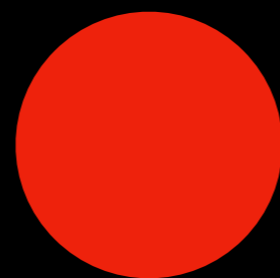
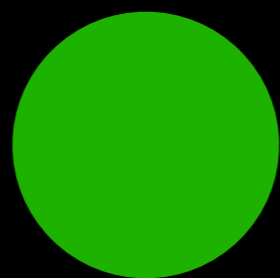
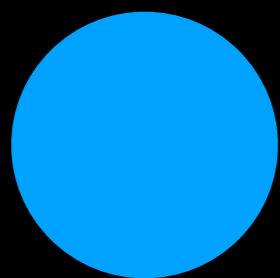


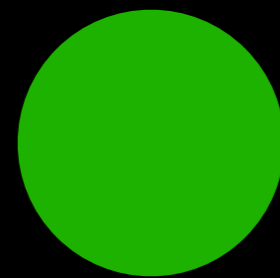
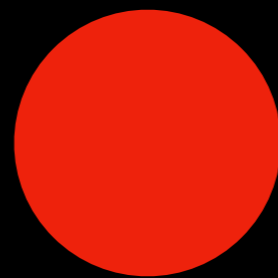
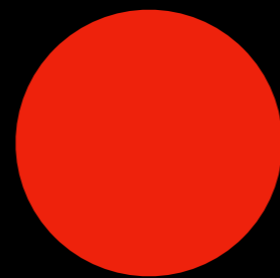
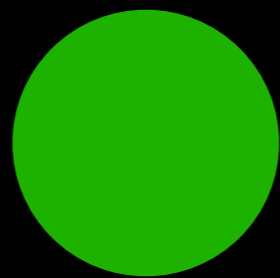
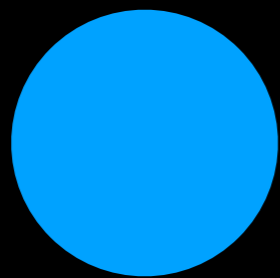


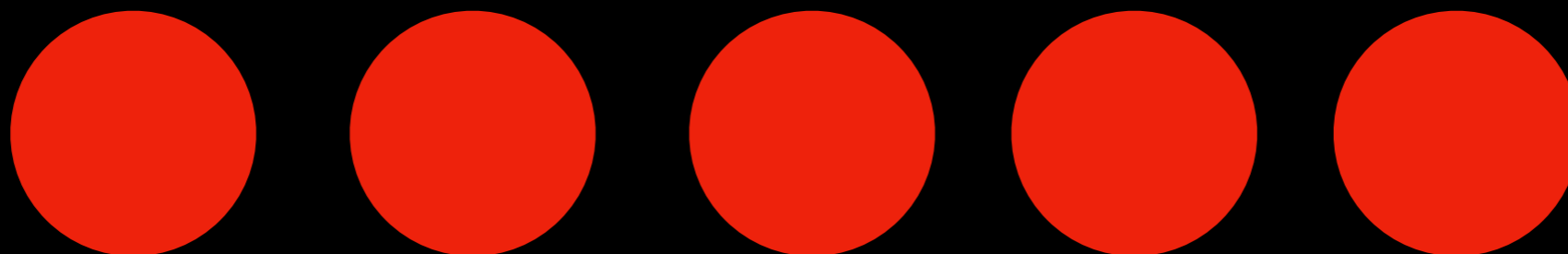
At most 2 red dots.

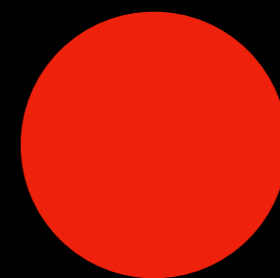
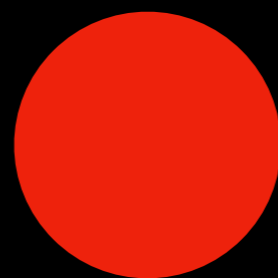
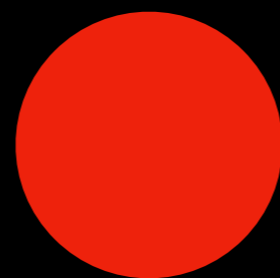
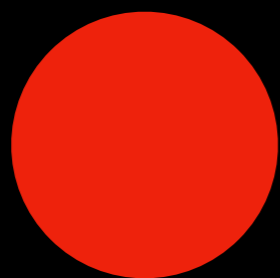


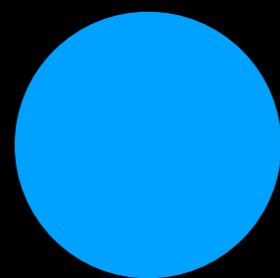
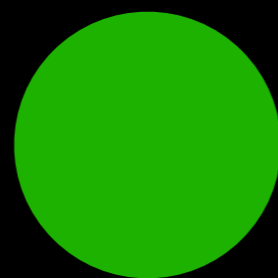
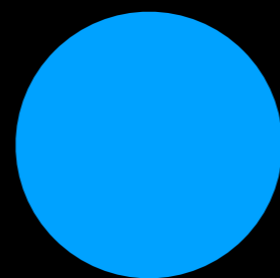
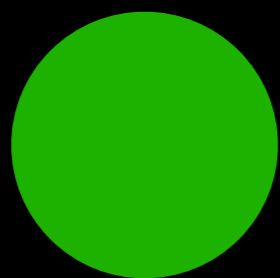


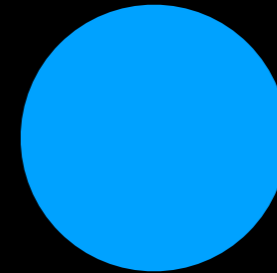
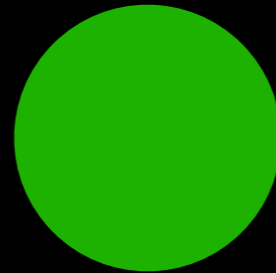
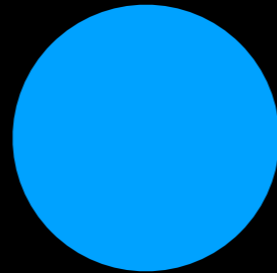
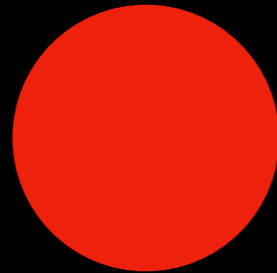
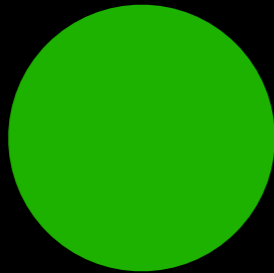


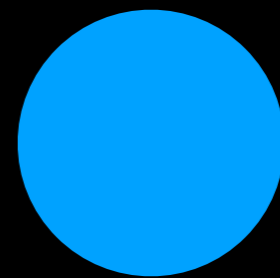
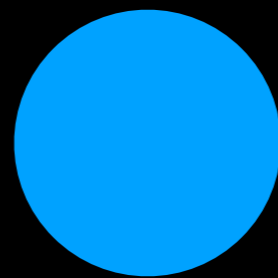
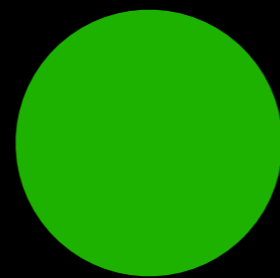
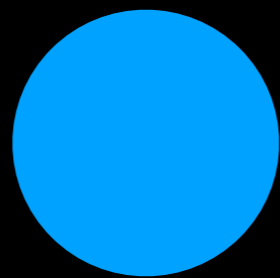
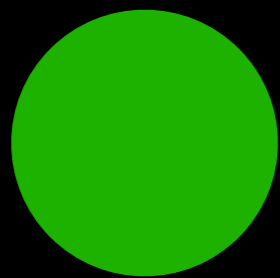


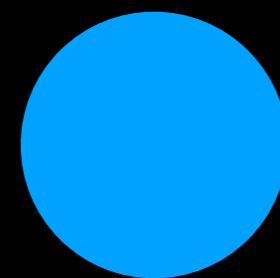
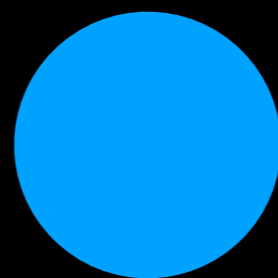
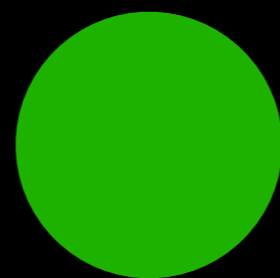
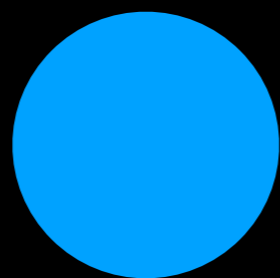
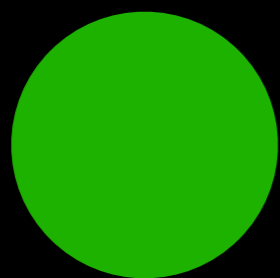


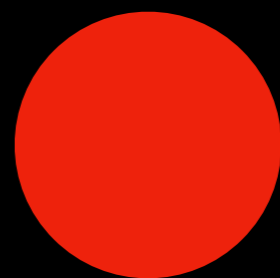
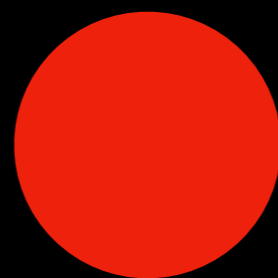
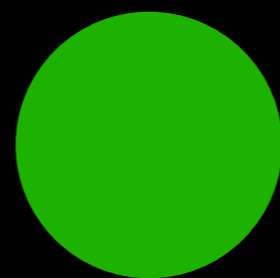
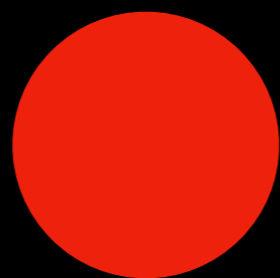
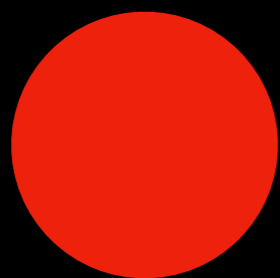


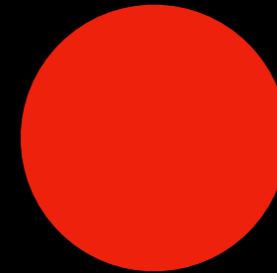
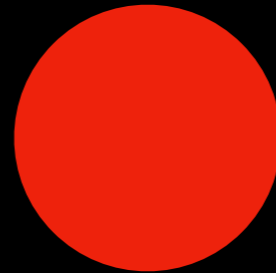
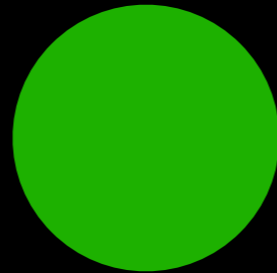
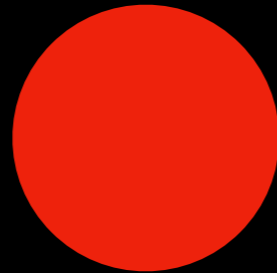
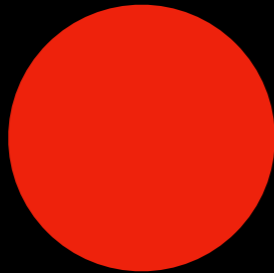














1, 2, or 4 red dots.

Learnability Prediction

monotone < connected < non-connected

Condition

Rules

Monotone

“There are 0, 1, or 2 red circles.”

“There are 3, 4, or 5 red circles.”

Connected

“There are 1, 2, or 3 red circles.”

“There are 2, 3, or 4 red circles.”

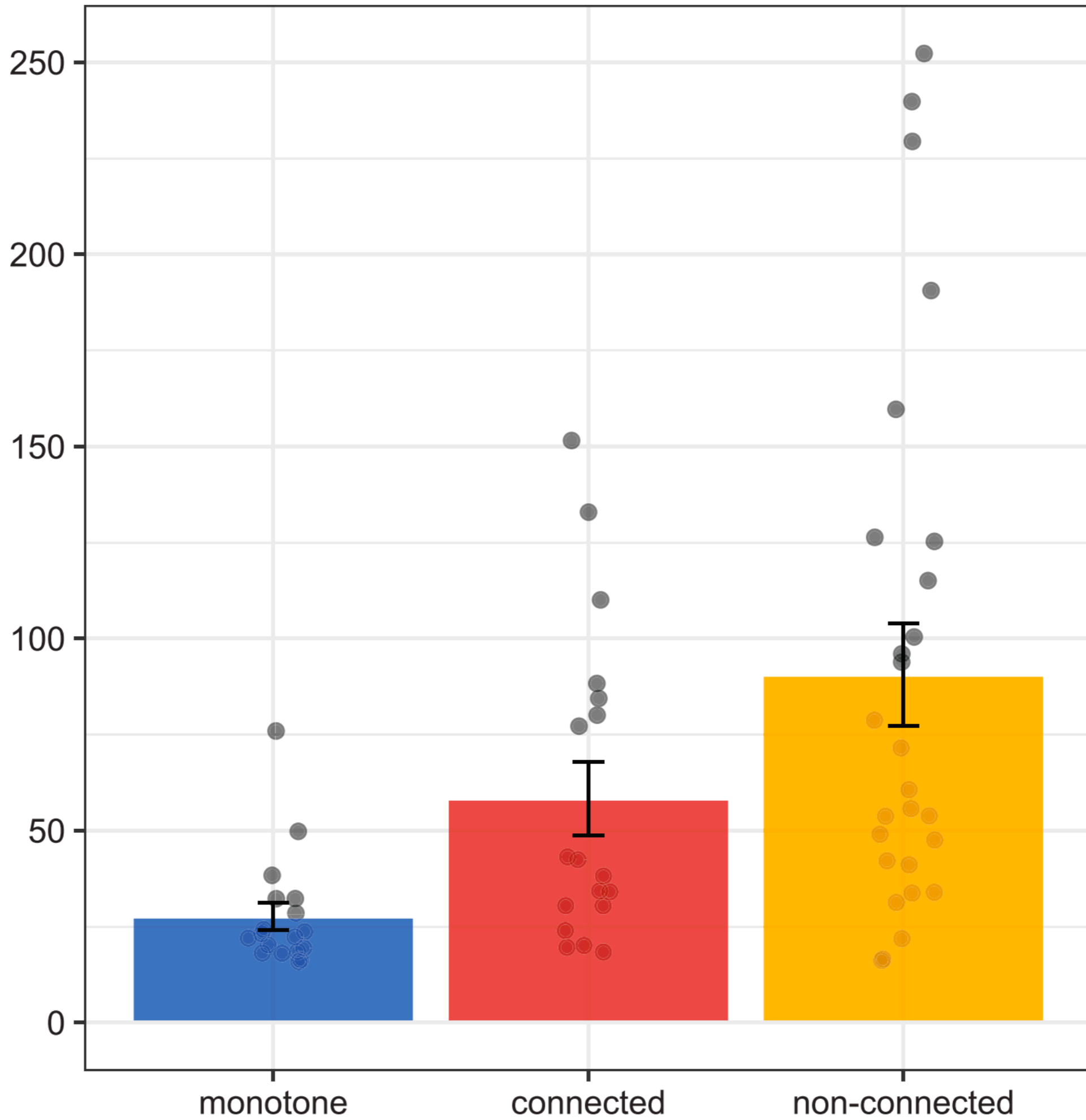
Non-connected

“There are 0, 1, or 5 red circles.”

“There are 0, 4, or 5 red circles.”

“There are 1, 2, or 4 red circles.”

Nb of Trials needed to reach the learning criterion



Subtleties

Subtleties

- Only monotone vs. non-connected was significant

Subtleties

- Only monotone vs. non-connected was significant
- “0, 4, or 5” and “0, 1, or 5” much faster than “1, 2, or 4”

Subtleties

- Only monotone vs. non-connected was significant
- “0, 4, or 5” and “0, 1, or 5” much faster than “1, 2, or 4”
- Connected re-coded as zero, one, or both (rule and negation) *is* then significantly different

Dynamic Analysis

- Odds of saying “yes” to n red dots, if already said “yes” to $n-1$ and $n+1$ in the same block significantly higher than if for only one or for neither.
- (even controlling for whether the actual rule is connected)

“The grand goal is to find a list of properties which are, in some sense, double universals: universals across languages, but also across word types....”

Constraints on the lexicons of human languages have cognitive roots present in baboons (*Papio papio*)

Emmanuel Chemla^a, Isabelle Dautriche^{b,c,d}, Brian Buccola^{a,1,2}, and Joël Fagot^{c,d}

^aLaboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, École Normale Supérieure, PSL University, École des Hautes Études en Sciences Sociales, CNRS, 75005 Paris, France; ^bCentre for Language Evolution, University of Edinburgh, Edinburgh EH8 9YL, United Kingdom; ^cLaboratory of Cognitive Psychology, Aix-Marseille University, CNRS, 13331 Marseille Cedex 3, France; and ^dInstitute of Language, Communication, and the Brain, Aix-Marseille University, CNRS, 13604 Aix-en-Provence Cedex 1, France

Edited by Barbara H. Partee, University of Massachusetts Amherst, Amherst, MA, and approved June 14, 2019 (received for review April 24, 2019)

Using a pattern extraction task, we show that baboons, like humans, have a learning bias that helps them discover connected patterns more easily than disconnected ones—i.e., they favor rules like “contains between 40% and 80% red” over rules like “contains around 30% red or 100% red.” The task was made as similar as possible to a task previously run on humans, which was argued to reveal a bias that is responsible for shaping the lexicons of human languages, both content words (nouns and adjectives) and logical words (quantifiers). The current baboon result thus suggests that the cognitive roots responsible for regularities across the content and logical lexicons of human languages are present in a similar form in other species.

connectedness | human languages and their lexicons | primate semantics

























Humans and animals categorize objects in the world into

show that humans have corresponding learning biases favoring connected quantifiers, as evidenced by performance on rule learning, or pattern extraction, tasks: It is easier to discover connected rules than nonconnected ones, and easier still to discover monotone ones.

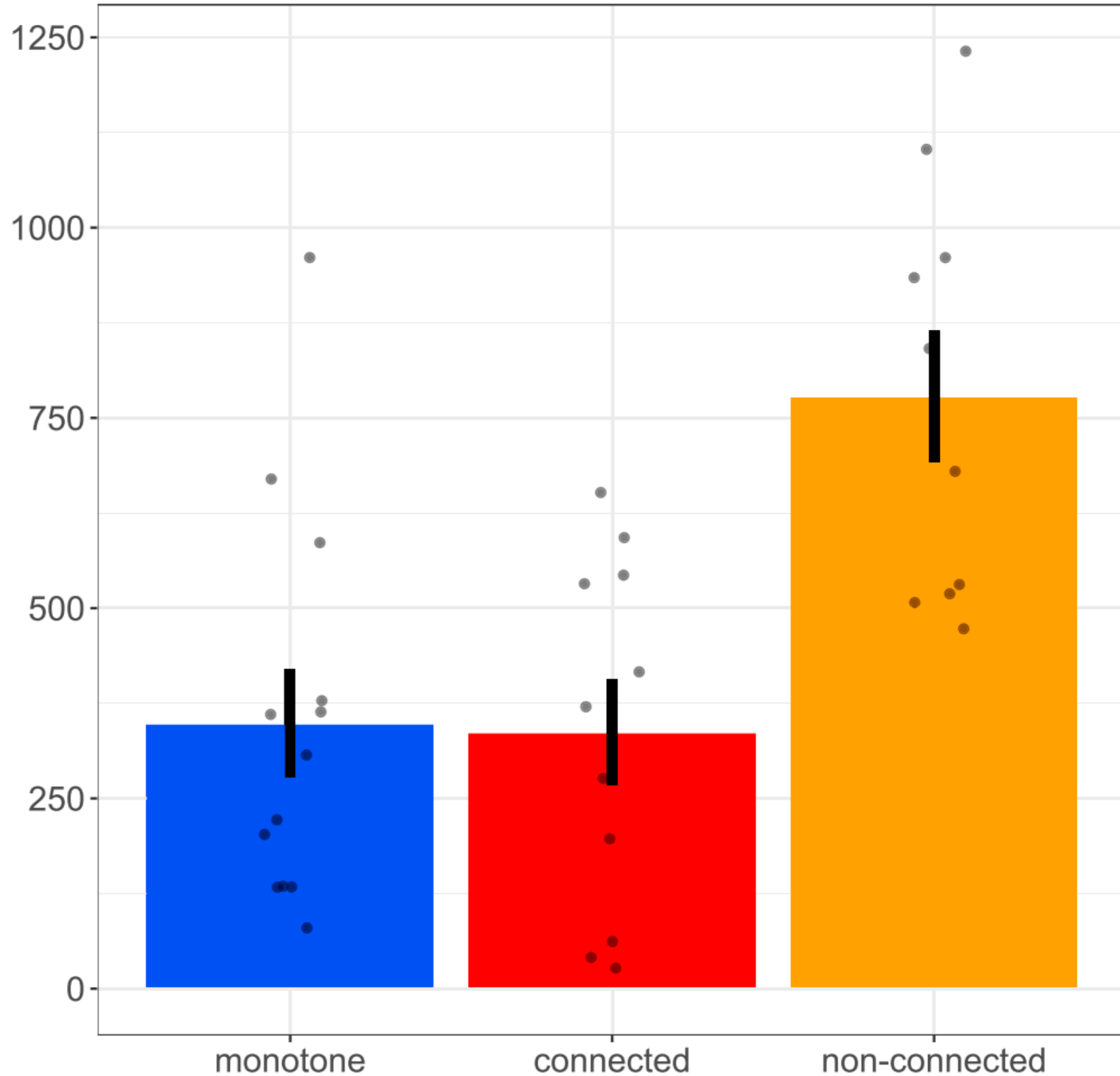
A natural hypothesis is that the source of the regularity of the world's lexicons, for both content and logical words, is a learning bias for connectedness. Indeed, Chemla et al. (4) argue that their experimental results with humans support this

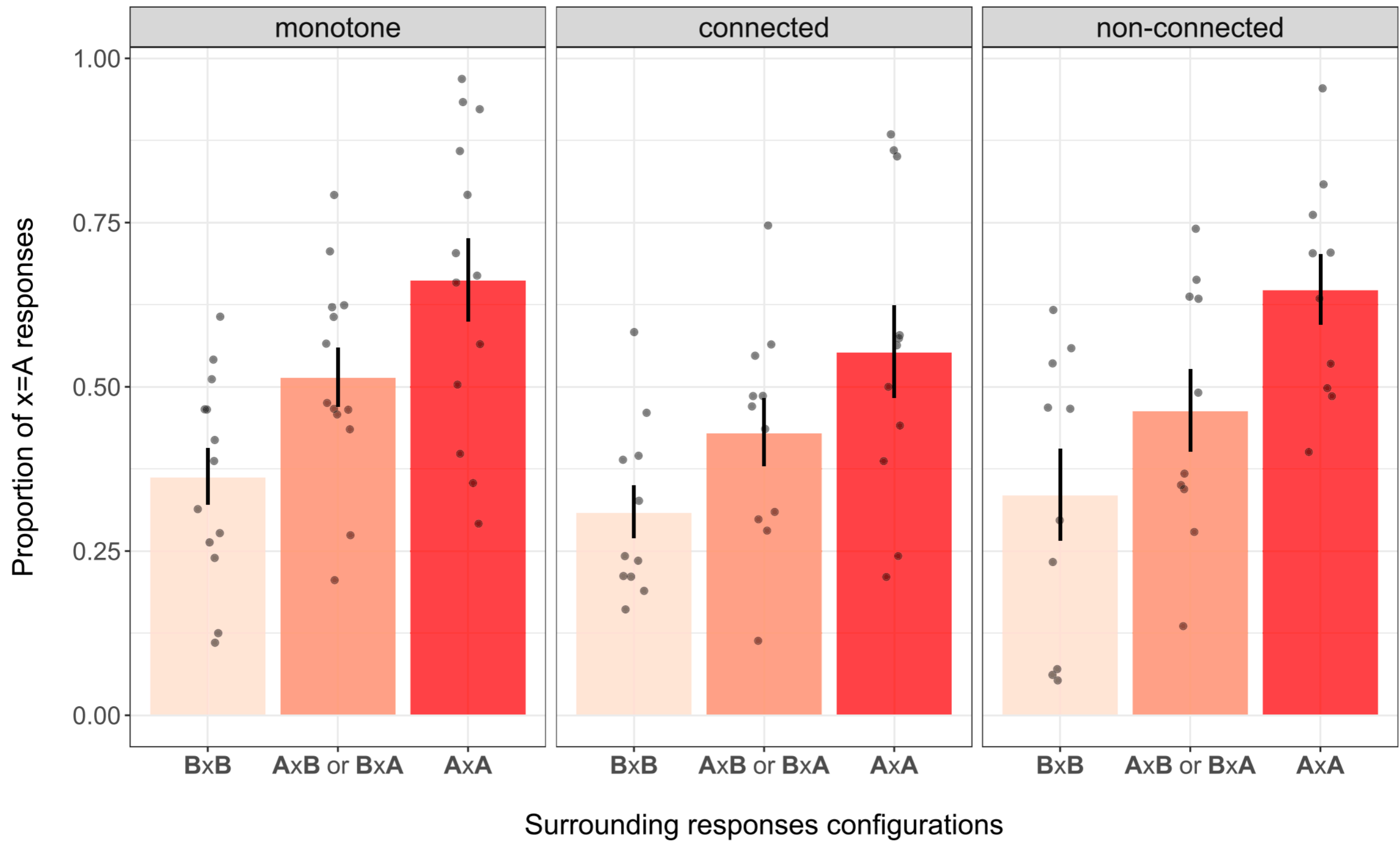
Significance

Universals in language are hard to come by, yet one candidate is that words across the lexicons of the world's languages are, by and large, connected: When a word applies to two

	Displays						Response buttons (comparison stimuli)	
	0%	20%	40%	60%	80%	100%		
Set 1								
Set 2								
Set 3								

Nb of blocks needed to reach the learning criterion





“The connectedness constraint is thus active in [humans and baboons] in a form that can explain how the referential and functional lexicons of human languages are shaped.”

Discussion

- Very small-scale
- No linguistic prompt (e.g. “gleeb of the dots are red”)
- Connectedness vs. monotonicity?

**What do we know about
quantifier acquisition?**

Some burning questions?

- What is the order of acquisition of quantifiers?
- Is it fixed, like that of number words, or does it vary?
- Which cognitive systems constrain it?
- Which logical properties constrain it?

Cross-linguistic patterns in the acquisition of quantifiers

Napoleon Katsos^{a,1}, Chris Cummins^b, Maria-José Ezeizabarrena^c, Anna Gavarró^d, Jelena Kuvač Kraljević^e, Gordana Hrzić^e, Kleanthes K. Grohmann^f, Athina Skordi^g, Kristine Jensen de López^h, Lone Sundahl^h, Angeliek van Houtⁱ, Bart Hollebrandseⁱ, Jessica Overwegⁱ, Myrthe Faber^{ij}, Margreet van Koert^k, Nafsika Smith^l, Maigi Vija^m, Sirli Zupping^m, Sari Kunnariⁿ, Tiffany Morisseau^o, Manana Rusieshvili^p, Kazuko Yatsushiro^q, Anja Fengler^r, Spyridoula Varlokosta^s, Katerina Konstantzou^s, Shira Farby^t, Maria Teresa Guasti^u, Mirta Vernice^u, Reiko Okabe^v, Miwa Isobe^w, Peter Crosthwaite^x, Yoonjee Hong^y, Ingrida Balčiūnienė^z, Yanti Marina Ahmad Nizar^g, Helen Grech^{aa}, Daniela Gatt^{aa}, Win Nee Cheong^{bb}, Arve Asbjørnsen^{cc}, Janne von Koss Torkildsen^{dd}, Ewa Haman^{ee}, Aneta Miękisz^{ee}, Natalia Gagarina^q, Julia Puzanova^{ff}, Darinka Anđelković^{gg}, Maja Savić^{gg}, Smiljana Jošić^{gg}, Daniela Slančová^{hh}, Svetlana Kapalkováⁱⁱ, Tania Barberán^c, Duygu Özge^{jj}, Saima Hassan^{kk}, Cecilia Yuet Hung Chan^{ll}, Tomoya Okubo^{mm}, Heather van der Lely^{nn,2}, Uli Sauerland^q, and Ira Noveck^o

^aDepartment of Theoretical and Applied Linguistics, University of Cambridge, Cambridge CB3 9DA, United Kingdom; ^bDepartment of Linguistics and English Language, University of Edinburgh, Edinburgh EH8 9AD, United Kingdom; ^cDepartment of Linguistics and Basque Studies, University of the Basque Country, ES-01006 Vitoria-Gasteiz, Spain; ^dDepartment of Catalan Philology, Universitat Autònoma de Barcelona, 08913 Bellaterra, Spain; ^eDepartment of Speech and Language Pathology, University of Zagreb, 10000 Zagreb, Croatia; ^fDepartment of English Studies, University of Cyprus, 1678 Nicosia, Cyprus; ^gIndependent researcher; ^hCenter for Developmental and Applied Psychological Science, Aalborg University, DK 9220 Aalborg East, Denmark; ⁱCenter for Language and Cognition Groningen, University of Groningen, 9700 AS Groningen, The Netherlands; ^jDepartment of Psychology, University of Notre Dame, Notre Dame, IN 46556; ^kAmsterdam Center for Language and Communication, University of Amsterdam, 1012 VT Amsterdam, The Netherlands; ^lHertfordshire Community National Health System Trust, Welwyn Garden City AL7 1BW, United Kingdom; ^mInstitute of Estonian and General Linguistics, University of Tartu, Tartu 50090, Estonia; ⁿFaculty of Humanities/Logopedics, University of Oulu, 90014 Oulu, Finland; ^oLaboratoire sur le Langage, le Cerveau, et la Cognition, CNRS-Université de Lyon, 69675 Bron Cedex, France; ^pDepartment of English Philology, Ivane Javakhishvili Tbilisi State University, 0179 Tbilisi, Georgia; ^qCentre for General Linguistics, D-10117 Berlin, Germany; ^rInstitute of Education, University of Leipzig, D-04109 Leipzig, Germany; ^sDepartment of Linguistics, University of Athens, 15784 Ilissia, Greece; ^tDepartment of English Literature and Linguistics, Bar Ilan University, Ramat-Gan 52900, Israel; ^uDepartment of Psychology, University of Milano-Bicocca, 20126 Milan, Italy; ^vCollege of Law, Nihon University, Tokyo 101-8375, Japan; ^wTraining Center for Foreign Languages and Diction, Tokyo University of the Arts, 110-8714 Tokyo, Japan; ^xCentre for Applied English Studies, University of Hong Kong, Hong Kong; ^yDepartment of Second Language Acquisition, University of Maryland, College Park, MD 20742; ^zFaculty of Humanities, Vytautas Magnus University, LT-44244 Kaunas, Lithuania; ^{aa}Department of Communication Therapy, University of Malta, Msida MSD 2080, Malta; ^{bb}Department of Psychology, HELP University, 50490 KL, Malaysia; ^{cc}Bergen Cognition and Learning Group, University of Bergen, 5009 Bergen, Norway; ^{dd}Faculty of Educational Sciences, University of Oslo, 0318 Oslo, Norway; ^{ee}Faculty of Psychology, University of Warsaw, 00-183 Warsaw, Poland; ^{ff}Institute of International Relations, Herzen State Pedagogical University of Russia, 191186 St. Petersburg, Russia; ^{gg}Laboratory for Experimental Psychology, University of Belgrade, Belgrade 11000, Serbia; ^{hh}Institute of Slovak Studies, General Linguistics and Media Studies, Presov University, 080 01 Presov, Slovakia; ⁱⁱDepartment of Speech Therapy, Comenius University, 818 06 Bratislava 16, Slovakia; ^{jj}Department of Psychology, Koç University, 34450 Istanbul, Turkey; ^{kk}International Islamic University/National University of Modern Languages, Islamabad 44000, Pakistan; ^{ll}Department of Linguistics and Translation, City University of Hong Kong, Hong Kong SAR; ^{mm}National Center for University Entrance Examinations, Tokyo 153-8501, Japan; and ⁿⁿDepartment of Psychology, Harvard University, Cambridge, MA 02138

Edited by Barbara H. Partee, University of Massachusetts at Amherst, Amherst, MA, and approved June 20, 2016 (received for review February 18, 2016)

Learners of most languages are faced with the task of acquiring words to talk about number and quantity. Much is known about the order of acquisition of number words as well as the cognitive and perceptual systems and cultural practices that shape it. Substantially less is known about the acquisition of quantifiers. Here we

These systems include an object-tracking system, which enables the precise representation of small quantities, and an analog magnitude system, which enables imprecise and approximate comparisons (1),

Expectations

- Monotonicity: up < down
- Totality: all, none < some, some...not, most
- Complexity: some < most
- Informativeness: truth > pragmatic felicity

Five Quantifiers

- 1+2: “all” >> “none”/“some” >> “some...not”
- 3: “some” >> “most”
- 4: false >> under-informative for “some,” “some...not,” and “most”

Participants

- 768 children (mean age = 5.5; range = 5.00–5.11; 398 female)
- 536 adults (293 female).
- 31 languages: Basque, Cantonese (Yue) Chinese, Catalan, Croatian, Cypriot Greek, Danish, Dutch, English, Estonian, Finnish, French, Georgian, German, Greek, Hebrew, Italian, Japanese, Korean, Lithuanian, Malay (Kuala Lumpur variety), Maltese, Mandarin Chinese, Norwegian, Polish, Russian, Serbian, Slovak, Spanish, Tamil, Turkish, and Urdu)
- 15 language genera (Baltic, Chinese, Finnic, Germanic, Greek, Indic, Japonic, Karto-Zan, Korean, Malayo-Sumbawan, Romance, Semitic, Slavic, Southern Dravidian, and Turkic)
- 11 language types [8 of the main language families in the world (Afro-Asiatic, Altaic, Austronesian, Dravidian, Indo-European, Kartvelian, Sino-Tibetan, and Uralic/Finno-Ugric] as well as 3 language isolates (Basque, Japonic, and Korean)

Procedure

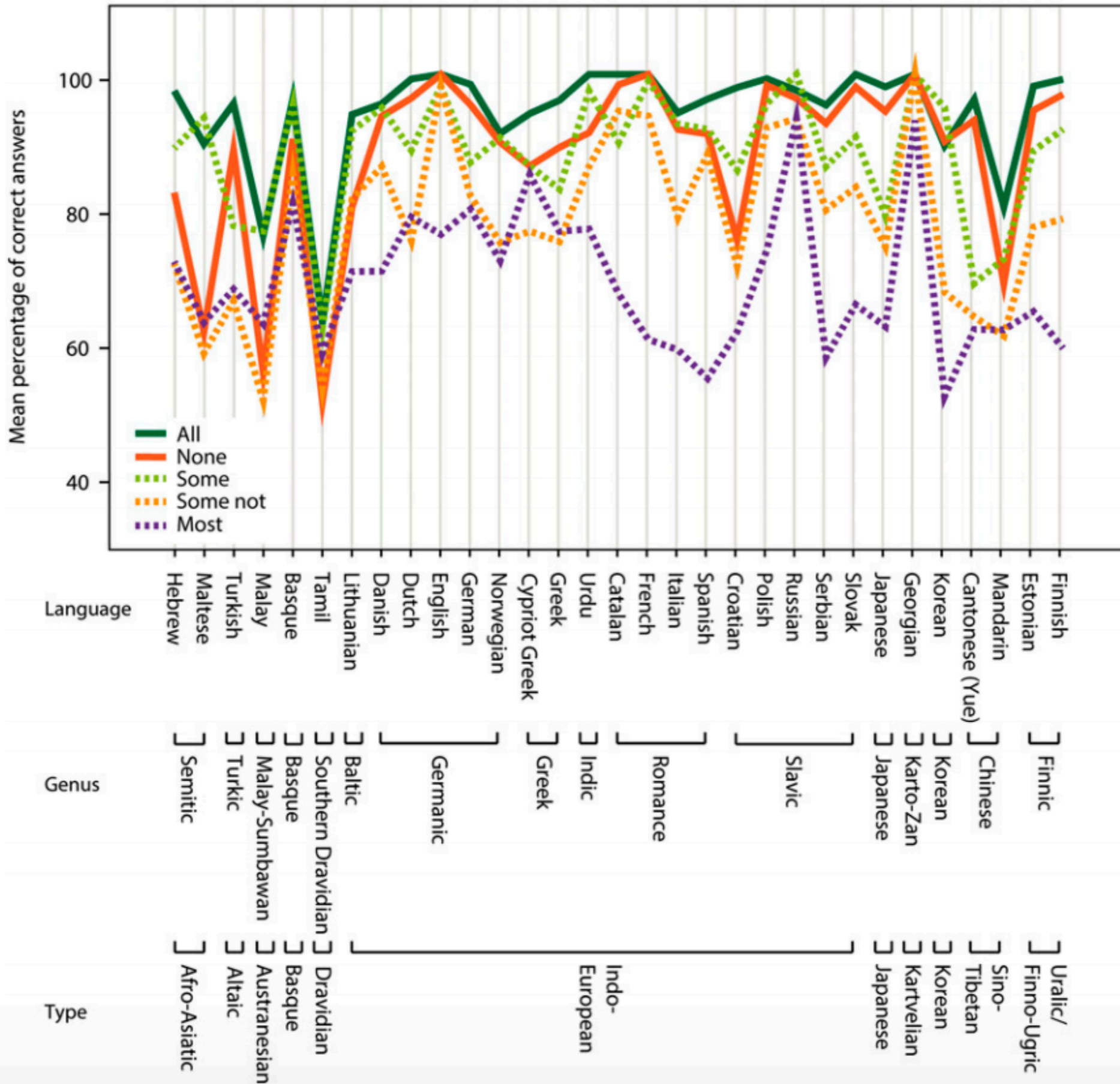


- Cave-girl task
- How many toys are in the boxes?
- [Quantifier] (of the) [objects] are (not) in the boxes.
- True or False (why?)
- True and informative/False/True but underinformative

Results

- Adults: 99% accuracy, 84% rejected underinformative sentences
- Children: 82% accuracy, 51% rejected underinformative sentences
- increasing >> decreasing in 27 of 31 languages (Catalan, English, Georgian, and Korean being exceptions)
- total >> partial in 25 of 31 languages (Georgian, Korean, Malay, Maltese, Russian, and Tamil being exceptions)
- some >> most in all 31 languages
- Children rejected underinformative uses less often in all 31 languages, adults in 28 languages (Cantonese, Russian, Urdu)

Percentage of correct responses for each quantifier by 5-year-old children



Discussion

- Order of acquisition of (only) five common quantifiers
- Hypothesis for potential cross-linguistic constraints
- But how to formalize those constraints, e.g., complexity?

Scalar quantifiers: Logic, acquisition, and processing

Bart Geurts

Department of Philosophy, University of Nijmegen, Nijmegen, the Netherlands

Napoleon Katsos and Chris Cummins

*Research Centre for English and Applied Linguistics, University of Cambridge,
Cambridge, UK*

Jonas Moons and Leo Noordman

Department of Psychology, University of Nijmegen, Nijmegen, the Netherlands

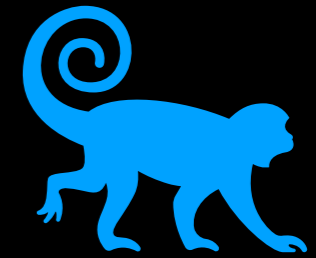
Superlative quantifiers ('at least 3', 'at most 3') and comparative quantifiers ('more than 2', 'fewer than 4') are traditionally taken to be interdefinable: the received view is that 'at least n ' and 'at most n ' are equivalent to 'more than $n-1$ ' and 'fewer than $n+1$ ', respectively. Notwithstanding the prima facie plausibility of this claim, Geurts and Nouwen (2007) argue that superlative quantifiers have essentially richer meanings than comparative ones. Geurts and Nouwen's theory makes three kinds of predictions that can be tested by experimental means. First, it predicts that superlative and comparative quantifiers should give rise to different patterns of reasoning. Second, the theory leads us to expect that children will master comparative quantifiers before superlative ones. Third, superlative quantifiers should be harder to process than comparative ones. We present three experiments that confirm these predictions.

Superlative vs. comparative quantifiers

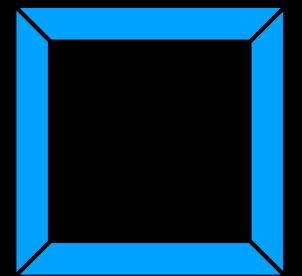
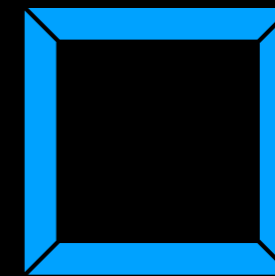
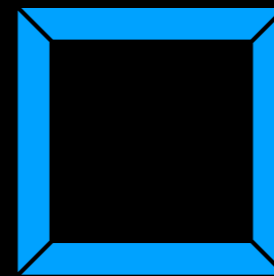
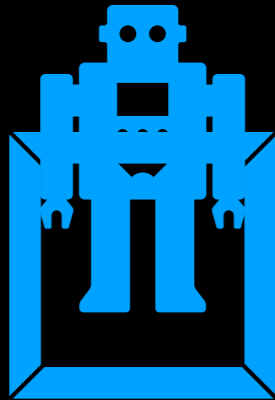
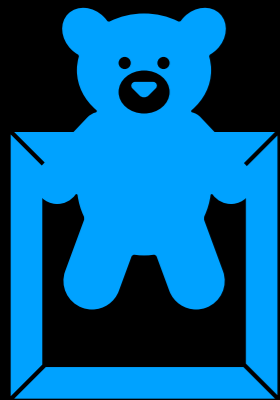
- “at least 3” = “more than 2”?
- “at most 3” = “fewer than 4”?
- Geurts and Nouwen (2007): No, superlatives are more complex
 1. **Berta didn't have at most 3 martinis.*
 2. *Berta didn't have fewer than 4 martinis.*

“At most n A are B ” means that the speaker – considers it possible that there is a set of n A 's that are B , and – is certain that there is no larger set of A 's that are B .

Experiment



- Q boxes have a toy



- Make the boxes and toys match the sentence

Participants

- 35 normally-developing 10- and 11-year-old children were recruited from a primary school (18 female, mean age 10.8, range 10.2-11.5)
- 35 adults were recruited from the student pool of the University of Cambridge (22 female, mean age 22.3, range 19.1-24.3).

Results

	2-arrangement		3-arrangement		4-arrangement		% all correct (sd)
	response	%	response	%	response	%	
exactly 3	0	0	0	100	0	0	100
	+1	100	+1	0	+1	0	
	+2	0	+2	0	+2	0	
	+3	0	+3	0	-1	100	
	+4	0	-1	0	-2	0	
	-1	0	-2	0	-3	0	
	-2	0	-3	0	-4	0	
% correct		100	100	100	100		
at least 3	0	11	0	71	0	46	88 (35)
	+1	57	+1	20	+1	6	
	+2	29	+2	3	+2	3	
	+3	3	+3	0	-1	37	
	+4	0	-1	6	-2	9	
	-1	0	-2	0	-3	0	
	-2	0	-3	0	-4	0	
% correct		89	94	91			
at most 3	0	23	0	46	0	29	43 (50)
	+1	40	+1	29	+1	23	
	+2	37	+2	11	+2	3	
	+3	0	+3	0	-1	37	
	+4	0	-1	14	-2	9	
	-1	0	-2	0	-3	0	
	-2	0	-3	0	-4	0	
% correct		63	60	46			
more than 3	0	0	0	0	0	80	97 (17)
	+1	3	+1	80	+1	17	
	+2	83	+2	17	+2	3	
	+3	11	+3	3	-1	0	
	+4	3	-1	0	-2	0	
	-1	0	-2	0	-3	0	
	-2	0	-3	0	-4	0	
% correct		97	100	100			
fewer than 3	0	80	0	6	0	3	77 (42)
	+1	0	+1	6	+1	0	
	+2	9	+2	6	+2	3	
	+3	0	+3	0	-1	14	
	+4	0	-1	83	-2	80	
	-1	11	-2	0	-3	0	
	-2	0	-3	0	-4	0	
% correct		91	83	80			

Discussion

- Intension vs. extension
- Is there an analogy with CONS? A need for stronger models/representations?
- What about frequency bias

**What do we know about
the distribution of
quantifiers in NL?**

Exploring the relation between semantic complexity and quantifier distribution in large corpora



Jakub Szymanik^{a,*}, Camilo Thorne^b

^a*Institute for Logic, Language and Computation, University of Amsterdam, P.O. Box 94242, 1090 GE Amsterdam, The Netherlands*

^b*Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Pfaffenwaldring 5b, D-70569 Stuttgart, Germany*

ARTICLE INFO

Article history:

Available online 7 March 2017

Keywords:

Generalized quantifiers

Semantic complexity

Corpus analysis

Generalized linear regression models

Analysis of deviance

ABSTRACT

In this paper we study if semantic complexity can influence the distribution of generalized quantifiers in a large English corpus derived from Wikipedia. We consider the minimal computational device recognizing a generalized quantifier as the core measure of its semantic complexity. We regard quantifiers that belong to three increasingly more complex classes: Aristotelian (recognizable by 2-state acyclic finite automata), counting ($k + 2$ -state finite automata), and proportional quantifiers (pushdown automata). Using regression analysis we show that semantic complexity is a statistically significant factor explaining 27.29% of frequency variation. We compare this impact to that of other known sources of complexity, both semantic (quantifier monotonicity and the comparative/superlative distinction) and superficial (e.g., the length of quantifier surface forms). In general, we observe that the more complex a quantifier, the less frequent it is.

© 2017 Elsevier Ltd. All rights reserved.

Corpus

- WaCky corpus (Baroni, 2009)
- Sentences ~ 43 millions
- Tokens ~ 800 milions
- POS-annotated

How to find quantifiers?

- E.g. word `most' can be:
 - a determiner (tag DT), or
 - adverb (RBR/RBS)
- When does it denote proportional quantifier?
- When followed by a plural noun (NNS) as in:
 - 'most/DT men/NNS',
- rather than an adjective (JJ) as in:
 - 'most/DT grateful/JJ.'

Linguistically plausible patterns

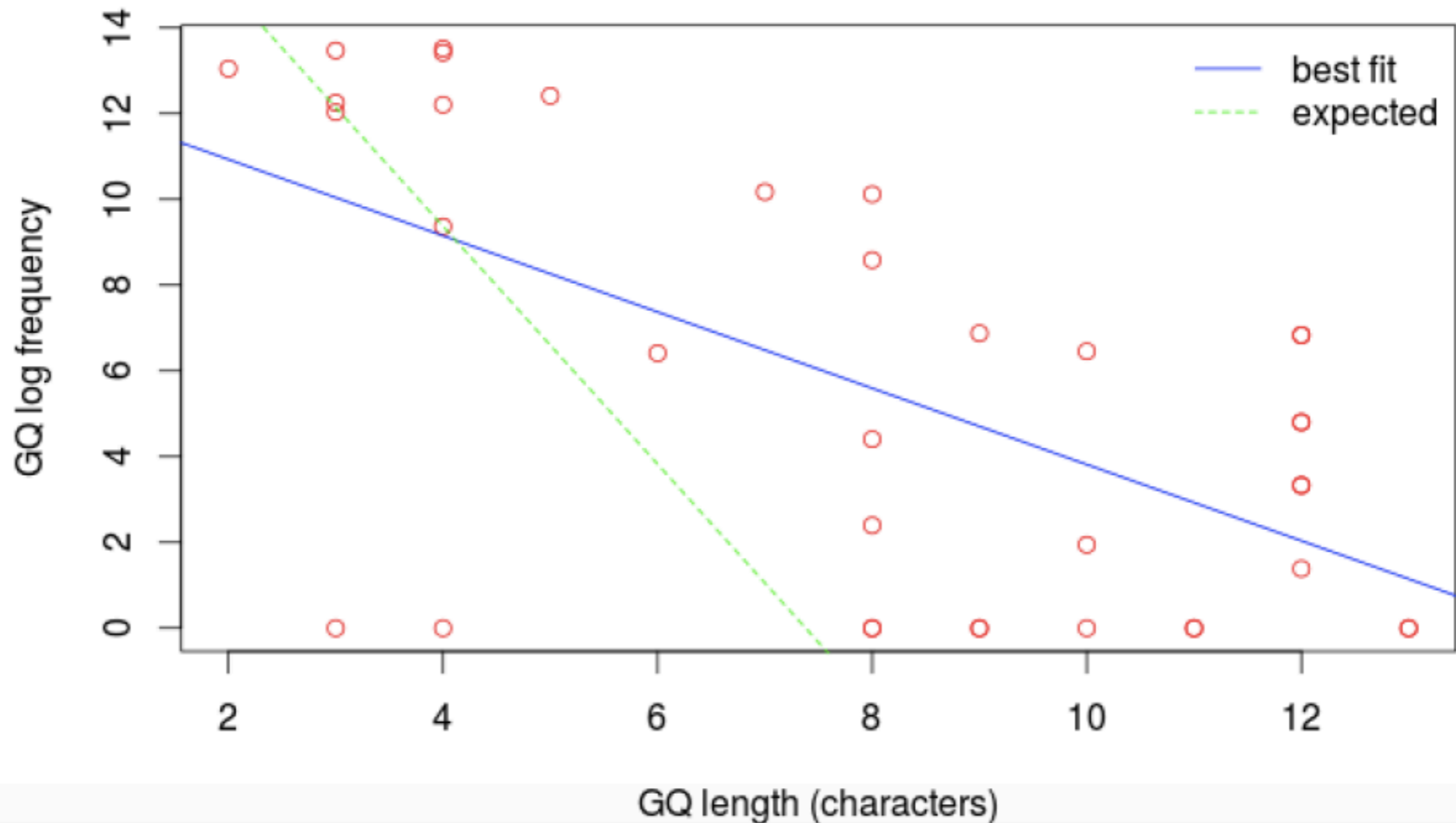
- E.g., GQ >k we matched with:
- *‘at/in least/jjs [a-z]{1,12}/cd’, viz., the preposition ‘at’ followed by the superlative adjective ‘least’ and a cardinal comprising up to 12 characters;*
- *‘more/rbr than/in [a-z]{1,12}/cd’, viz., the comparative adverb ‘more’ followed by the preposition ‘than’ and a cardinal;*
- *‘more/jjr than/in [a-z]{1,12}/cd’, viz., the same as before, but with ‘more’ a comparative adjective.*
- In total we counted occurrences of 36 patterns.

Descriptive analysis

- Aristotelian > counting > proportional, but also:
- short > multiword
- Both syntax and semantics influence.

Complexity and frequency

GQ distribution w.r.t. length (characters)



Factors influencing frequency



- Complexity: Aristotelian, counting, proportional.
- Monotonicity: 'up', 'down', and 'none'.
- Type: comparative or superlative.
- Length in words: number of word tokens.
- Length in characters: number of characters.



GLM regression analysis

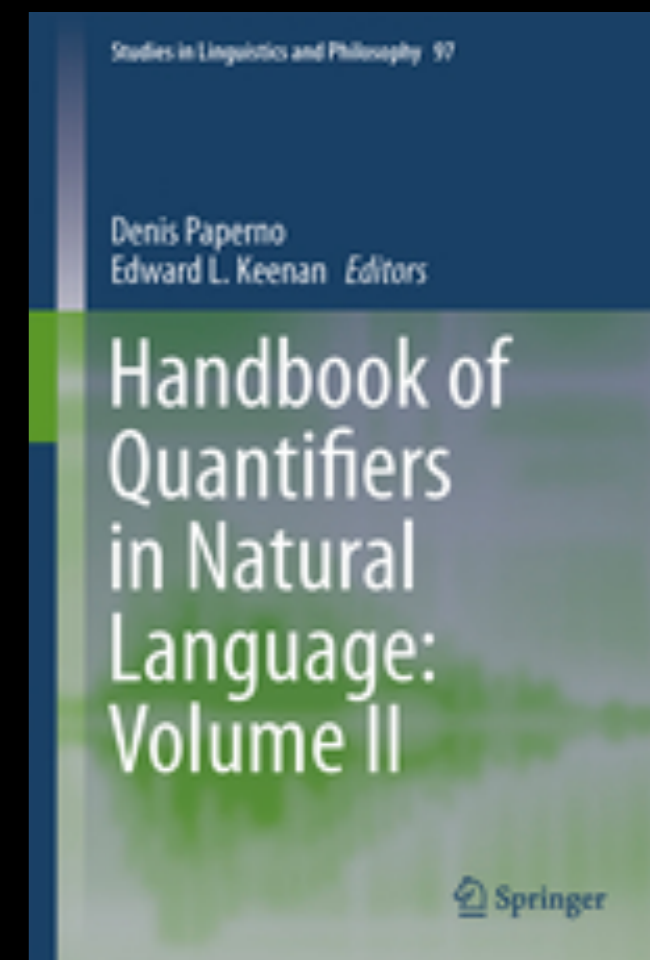
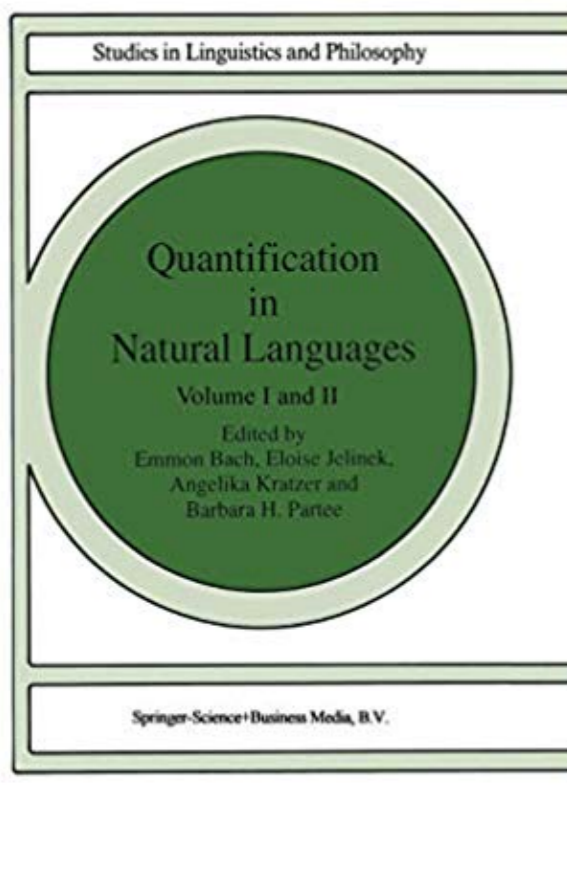
- GQ frequency as a complex function of various factors:
 - complexity (~27%),
 - type,
 - length (~47%),
 - right monotonicity (~26%).

Semantic complexity underlies GQ distributions

The Konstanz Universal Archive

Number	1199 (used to be 1203 in the old version)
Original	Every natural language has syntactic constituents (called noun-phrases) whose semantic function is to express generalized quantifiers over the domain of discourse.
Standardized	Every natural language has syntactic constituents (called noun-phrases) whose semantic function is to express generalized quantifiers over the domain of discourse.
Formula	
Keywords	quantification, generalized quantifier, NP
Domain	syntax, semantics
Type	unconditional
Status	achronic
Quality	absolute
Basis	unspecified
Source	Barwise & Cooper 1981 : 177, U1
Counterexamples	<p>By Frans Plank  03.08.2006, 09:49</p> <p>All the following lack NP quantification: Straits (Salish), Asurini (Tupi), Mohawk (Iroquoian), Lakhota (Siouan), Navajo (Athabaskan), Warlpiri (Pama-Nyungan), Gun-djeyhmi (Gunwingguan, Australian). Warlpiri and Gun-djeyhmi, for example, make use of verbal affixes to express various kinds of quantificational meaning. Asurini quantifiers such as <i>all</i>, <i>many</i>, <i>two</i> do not form a syntactic constituent with the noun because they do not belong to the category of determiners. They are members of other categories such as adverb, verb and noun instead. See discussion in Bach et al. 1995.</p>
Comments	<p>By Frans Plank  03.08.2006, 09:49</p> <p>According to Barwise & Cooper, it would "probably be wrong to claim that NPs are the only quantifiers in natural language. [...] It does seem reasonable, however, to claim that the noun-phrases of a language are all and only the quantifiers over the domain of discourse [...]."</p> <p>However, the universality of an NP constituent is dubious, pace Barwise & Cooper, for whom NPs typically host proper names and expressions accompanied by determiners such as 'every', 'most', 'one'. See further Bach et al.1995.</p>

Number	1203 (used to be 1207 in the old version)	
Original	Monotonicity constraint: The simple NPs of any natural language express monotone quantifiers or conjunctions of monotone quantifiers.	
Standardized	Monotonicity constraint: The simple NPs of any natural language express monotone quantifiers or conjunctions of monotone quantifiers.	
Formula		
Keywords	quantification, monotonicity, NP	
Domain	syntax, semantics	
Type	unconditional	
Status	achronic	
Quality	absolute	
Basis	unspecified	
Source	Barwise & Cooper 1981 : 187, U6	
Counterexamples	By Frans Plank 	03.08.2006, 09:49
Comments	By Frans Plank 	03.08.2006, 09:49
	Barwise & Cooper explain that this "proposed universal has the effect of ruling out many logically possible quantifiers as simple NP denotations. Examples are the denotations of 'an even number of men', 'exactly three or exactly five men', 'all but one man'. It seems unlikely that any natural language would have a basic determiner meaning 'an even number of', 'exactly three or exactly five' or 'all but one'."	



So what do we know?

- Keenan & Paperno generalize over 36 languages
- All Ls have proportional quantifiers
- All Ls have at least one lexical ONLY.
- 35 of 36 Ls present downward entailing DNPs
- Neither BJKP nor KP don't systematically review 'lexical universals'

Summary of the class

- 1: mathematical theory of quantification in natural language, with universal constraints
- 2: GQ theory enables natural computational representations
- 3: 'older' tools in formal learning do not explain universals
- 4: neural learnability explains many universals, across domains
- 5: the developmental and cross-linguistic pictures are not very complete

Discussion Time!

- Scaling up experiments:
 - Computational
 - Human
- Controlling semantic factors
- Unification of / more fundamental explanations of universals?
- Cross-linguistic work
- Complexity vs learnability

Thanks!

<http://jakubszymanik.com>
jakub.szymanik@gmail.com

<http://shane.st>
shanest@uw.edu