Learnability of Quantifiers

Developmental psychology and unification

LANGUAGE in INTERACTION

CLMBR

UWNLP

Recap

Recap

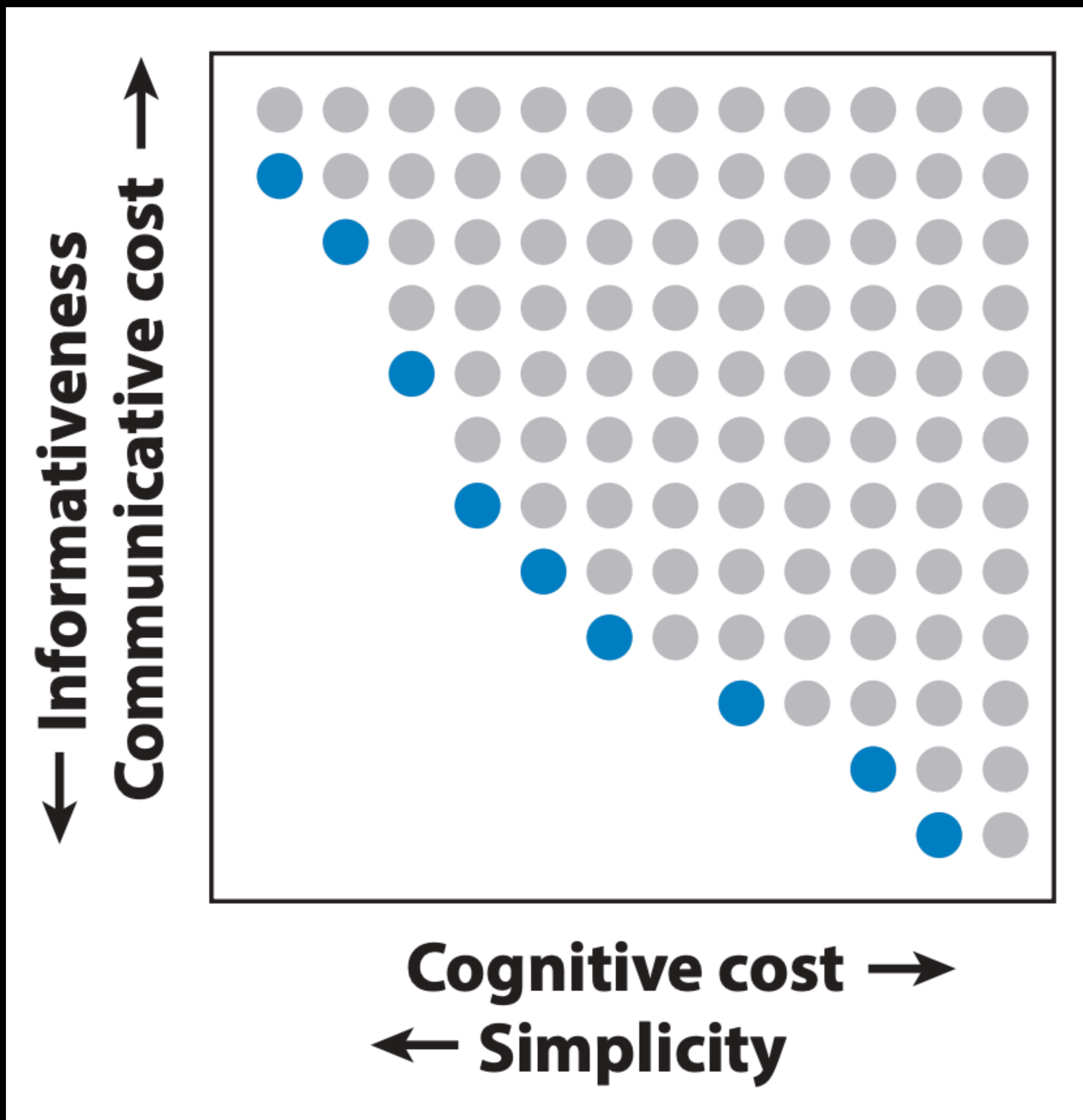
- Yesterday:
  - Ease of learning explains universals for quantifiers, responsive predicates, and color terms

Recap

- Yesterday:
  - Ease of learning explains universals for quantifiers, responsive predicates, and color terms
- Today:
  - Coda: efficient communication
  - (developmental) psychology of quantifiers
  - Cross-linguistic picture
  - Discussion

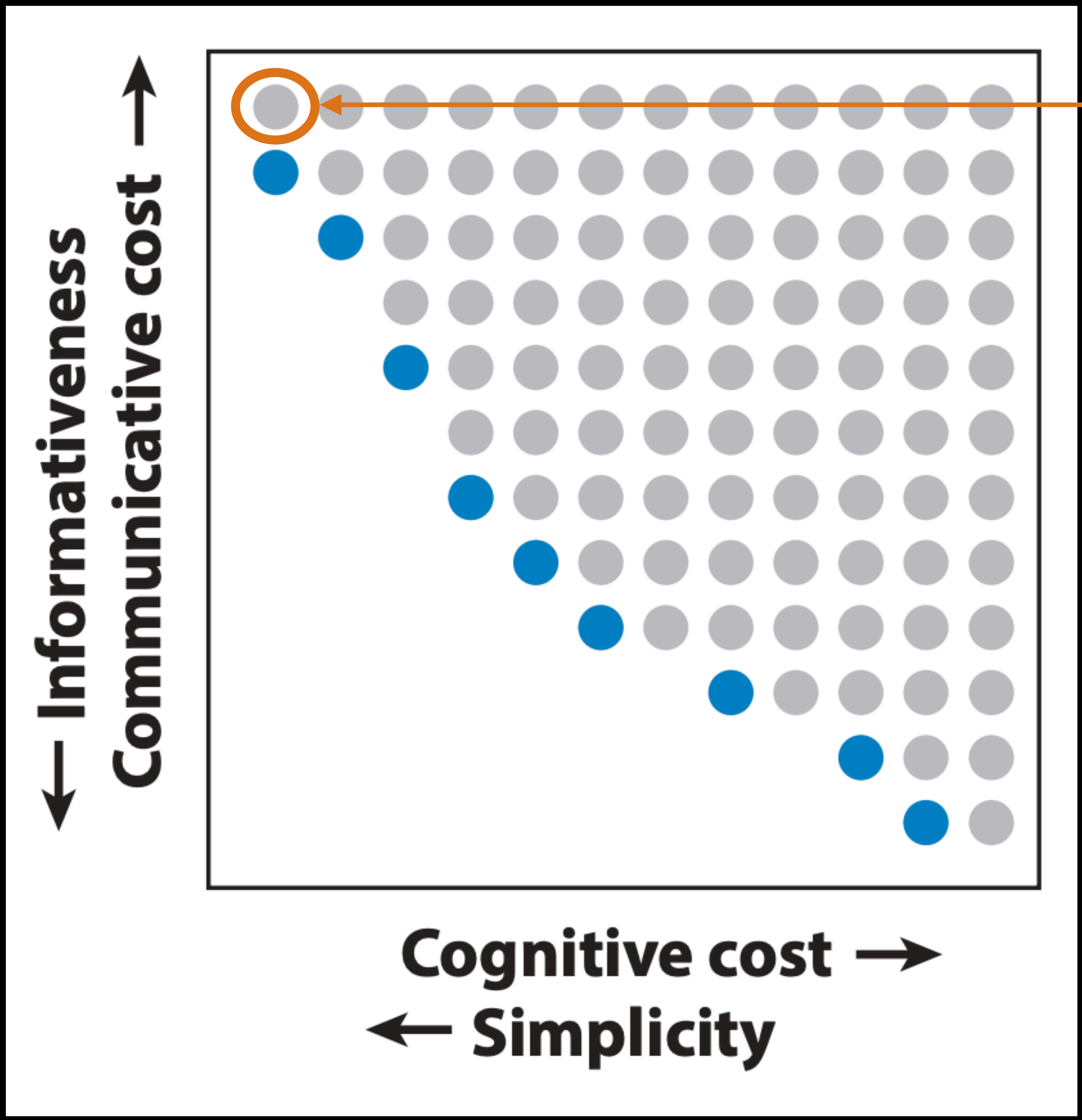
Efficient Communication

Simplicity and Informativeness Trade-off



Kemp, Xu, Regier 2018

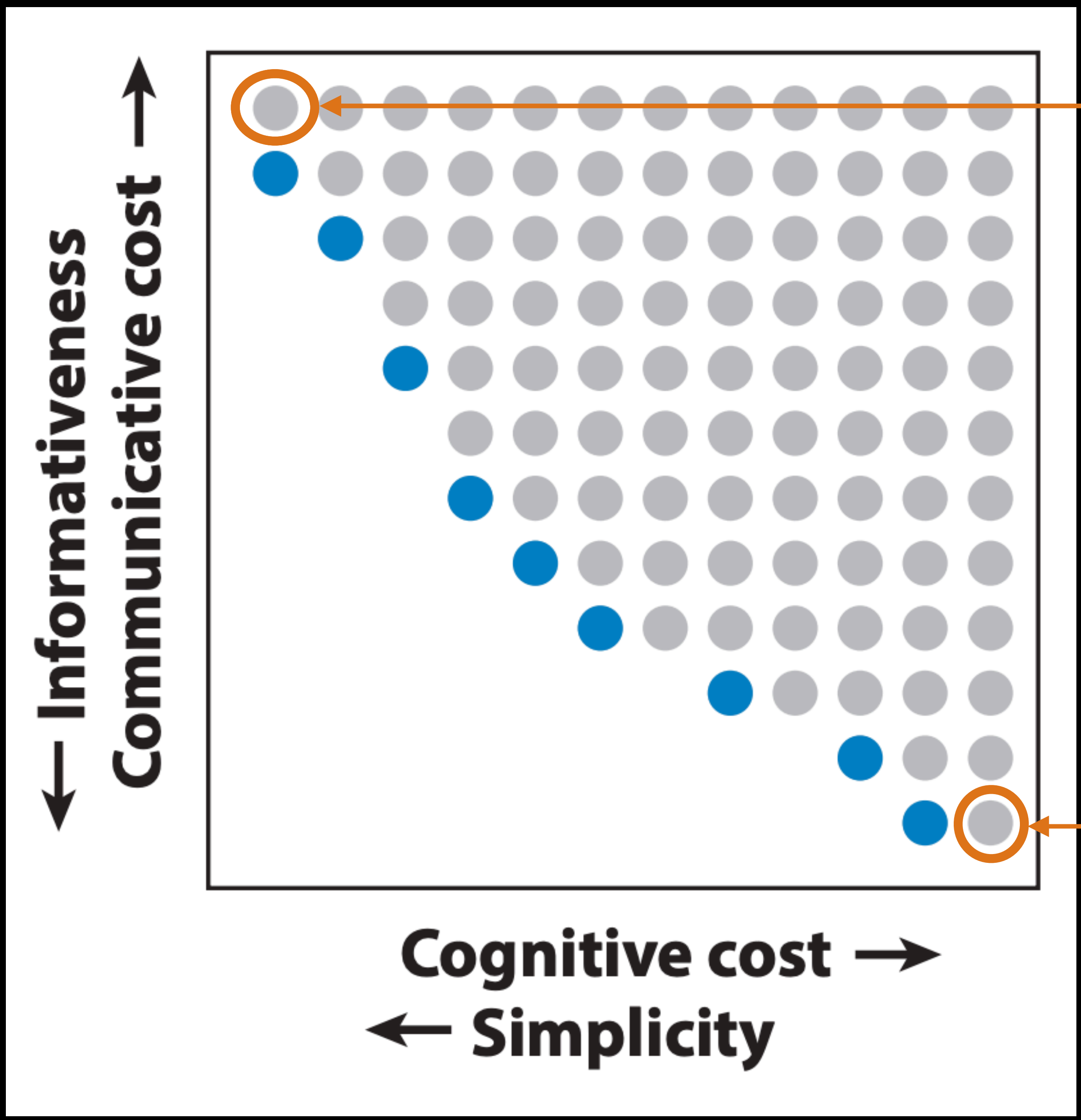
Simplicity and Informativeness Trade-off



L: {[[e]] = M}
Very simple
Very uninformative

Kemp, Xu, Regier 2018

Simplicity and Informativeness Trade-off

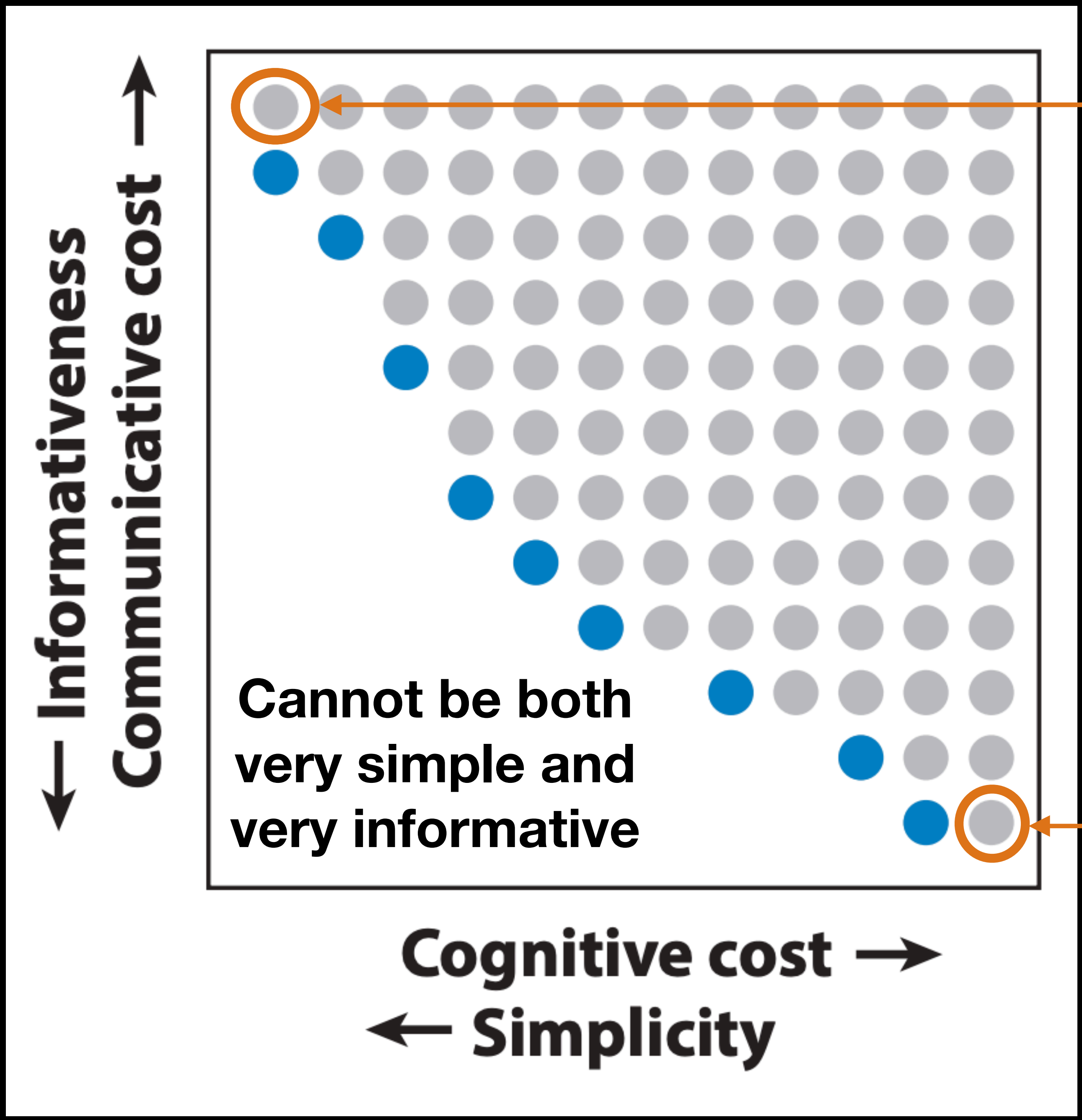


L: {[[e]] = M}
Very simple
Very uninformative

L: {[[e1]] = {m1}, [[e2]] = {m2}, …}
Very complex
Very informative

Kemp, Xu, Regier 2018

Simplicity and Informativeness Trade-off

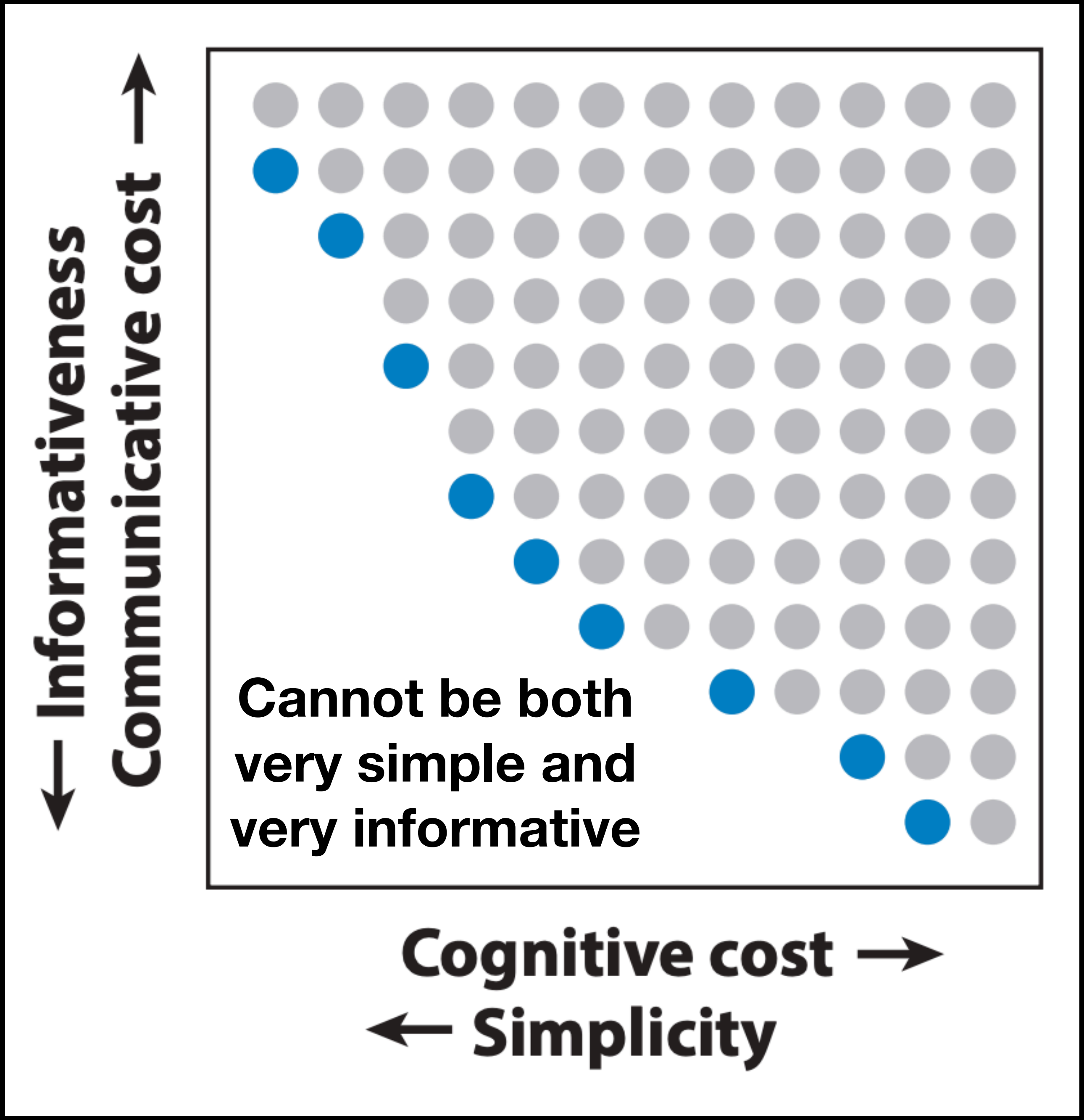


L: {[[e]] = M}
Very simple
Very uninformative

L: {[[e1]] = {m1}, [[e2]] = {m2}, …}
Very complex
Very informative

Kemp, Xu, Regier 2018

Simplicity and Informativeness Trade-off



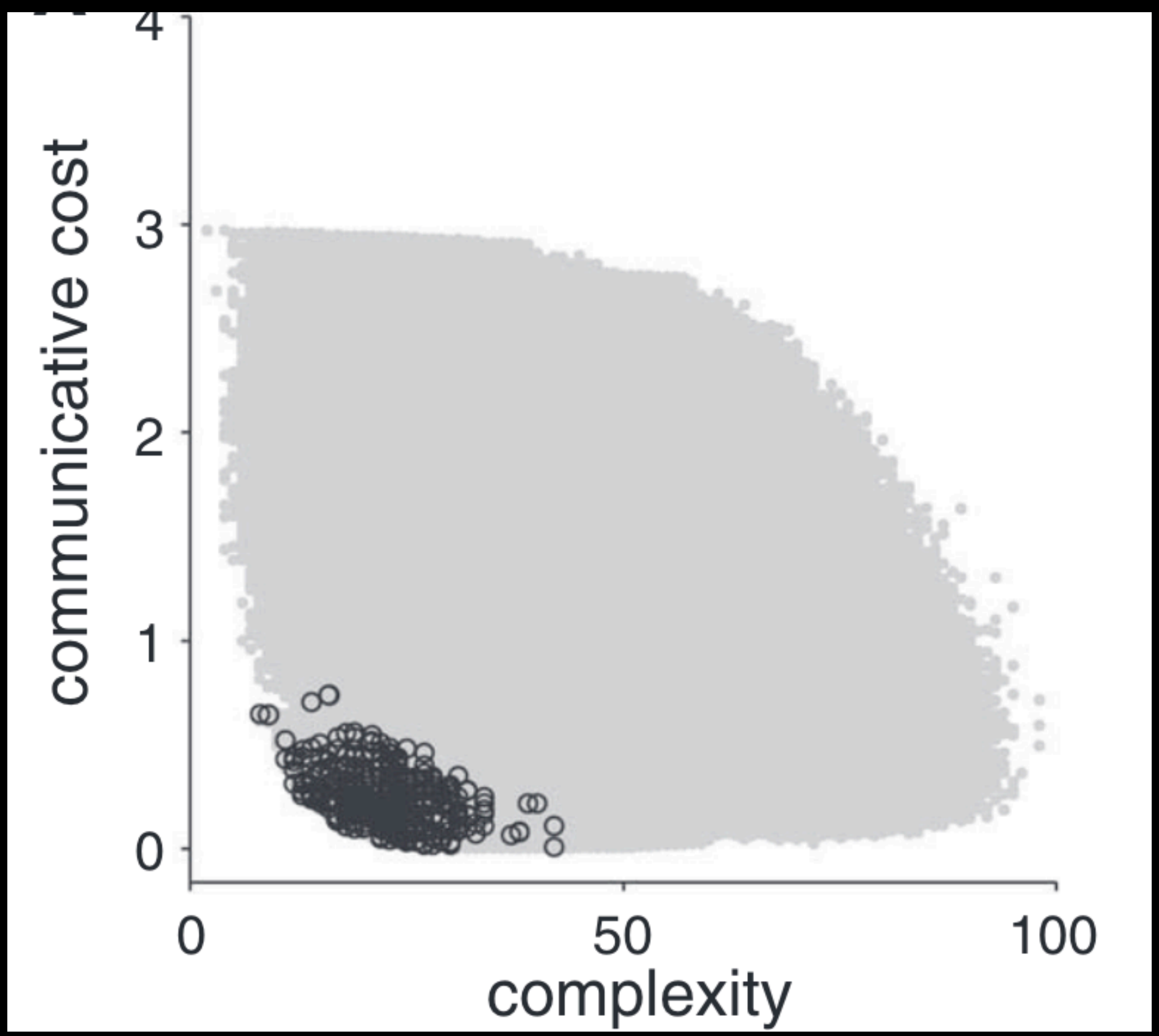
Pareto frontier: those languages doing *as well as possible*. (No other language is both simpler and more informative.)

Hypothesis: natural languages lie at (or near) the Pareto frontier.

Kemp, Xu, Regier 2018

Simplicity and Informativeness Trade-off

- Kinship terms (Kemp and Regier 2012)
- Color (Zaslavsky et al 2018)
- Person (Zaslavsky et al 2021)
- Quantifiers (Steinert-Threlkeld 2019, 2021)
- Indefinites (Denić, S-T, Szymanik 2020, 2022)
- Tense and evidentials (Mollica et al 2021)
- Logical vocabulary (Uegaki 2022)
- Modals (Imel, Steinert-Threlkeld 2022)
- …



Kemp and Regier 2012

entropy

Article

Quantifiers in Natural Language: Efficient Communication and Degrees of Semantic Universals

Shane Steinert-Threlkeld

Department of Linguistics, University of Washington, Seattle, WA 98195, USA; shanest@uw.edu

Abstract: While the languages of the world vary greatly, they exhibit systematic patterns, as well. Semantic universals are restrictions on the variation in meaning exhibit cross-linguistically (e.g., that, in all languages, expressions of a certain type can only denote meanings with a certain special property). This paper pursues an efficient communication analysis to explain the presence of semantic universals in a domain of function words: quantifiers. Two experiments measure how well languages do in optimally trading off between competing pressures of simplicity and informativeness. First, we show that artificial languages which more closely resemble natural languages are more optimal. Then, we introduce information-theoretic measures of degrees of semantic universals and show that these are not correlated with optimality in a random sample of artificial languages. These results suggest both that efficient communication shapes semantic typology in both content and function word domains, as well as that semantic universals may not stand in need of independent explanation.

Experiment

- Generate large number of artificial languages
- Simplicity: minimal expression length in LoT (cf day 3)
- Informativeness: average ability to successfully communicate an intended model
- Degree of naturalness: proportion of Qs that are (i) generalized universal, (ii) generalized existential, (iii) proportional
- Optimality: closeness to Pareto frontier

Informativeness

$$C(S, R) := 1 - I(S, R)$$

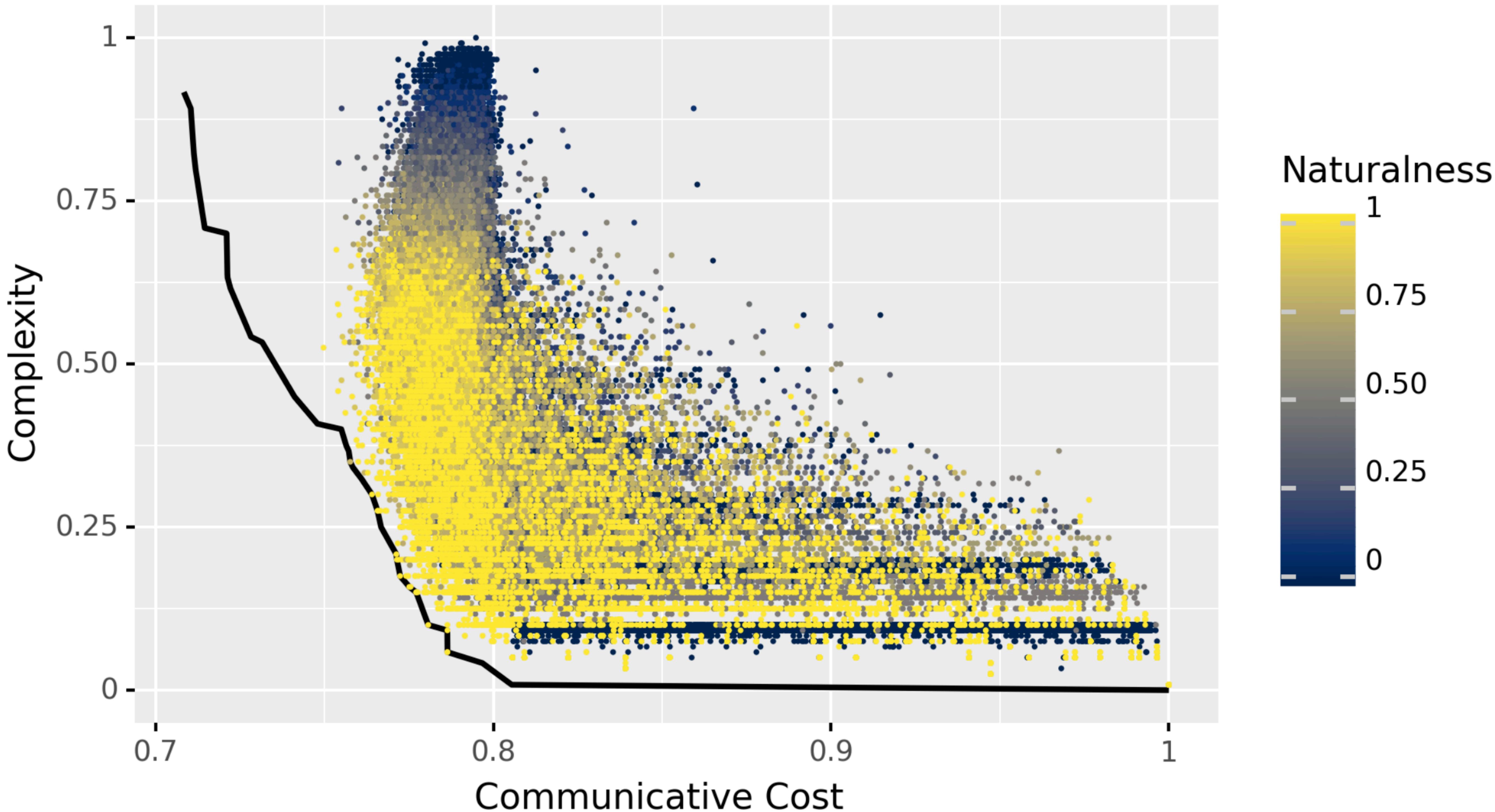
$$I(S, R) := \mathbb{E}_P[u(\mathbb{M}, \mathbb{M}')]$$

$$= \sum_{\mathbb{M}} P(\mathbb{M}) \sum_{Q} P_S(Q \mid \mathbb{M}) \sum_{\mathbb{M}'} P_R(\mathbb{M}' \mid Q) \cdot u(\mathbb{M}, \mathbb{M}')$$

Informativeness

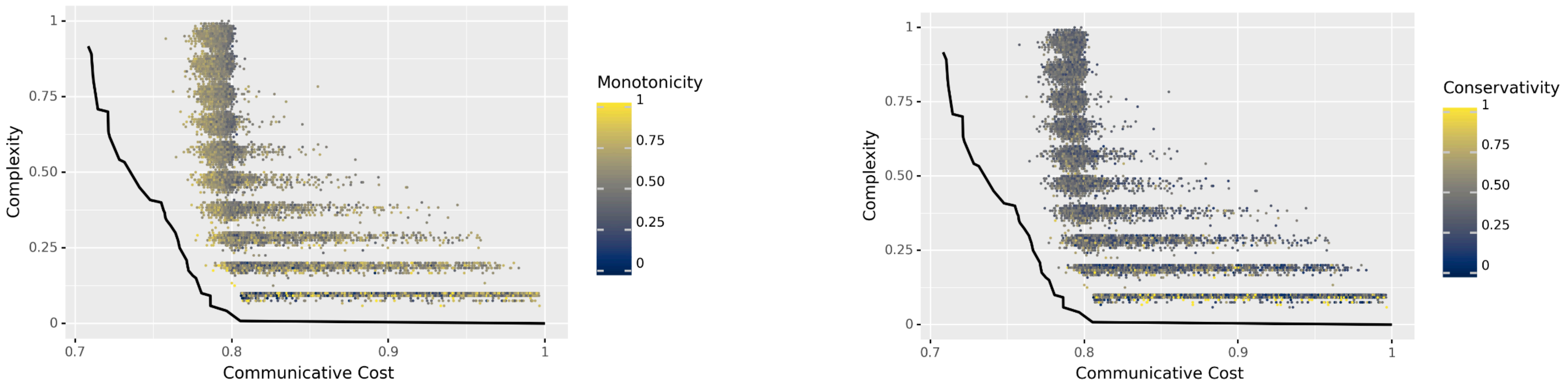
$$
\begin{aligned}
C(S,R) &:= 1 - I(S,R) \\
I(S,R) &:= \mathbb{E}_P[u(\mathbb{M}, \mathbb{M}')] \\
&= \sum_{\mathbb{M}} P(\mathbb{M}) \sum_{Q} P_S(Q \mid \mathbb{M}) \sum_{\mathbb{M}'} P_R(\mathbb{M}' \mid Q) \cdot u(\mathbb{M}, \mathbb{M}')
\end{aligned}
$$

Results



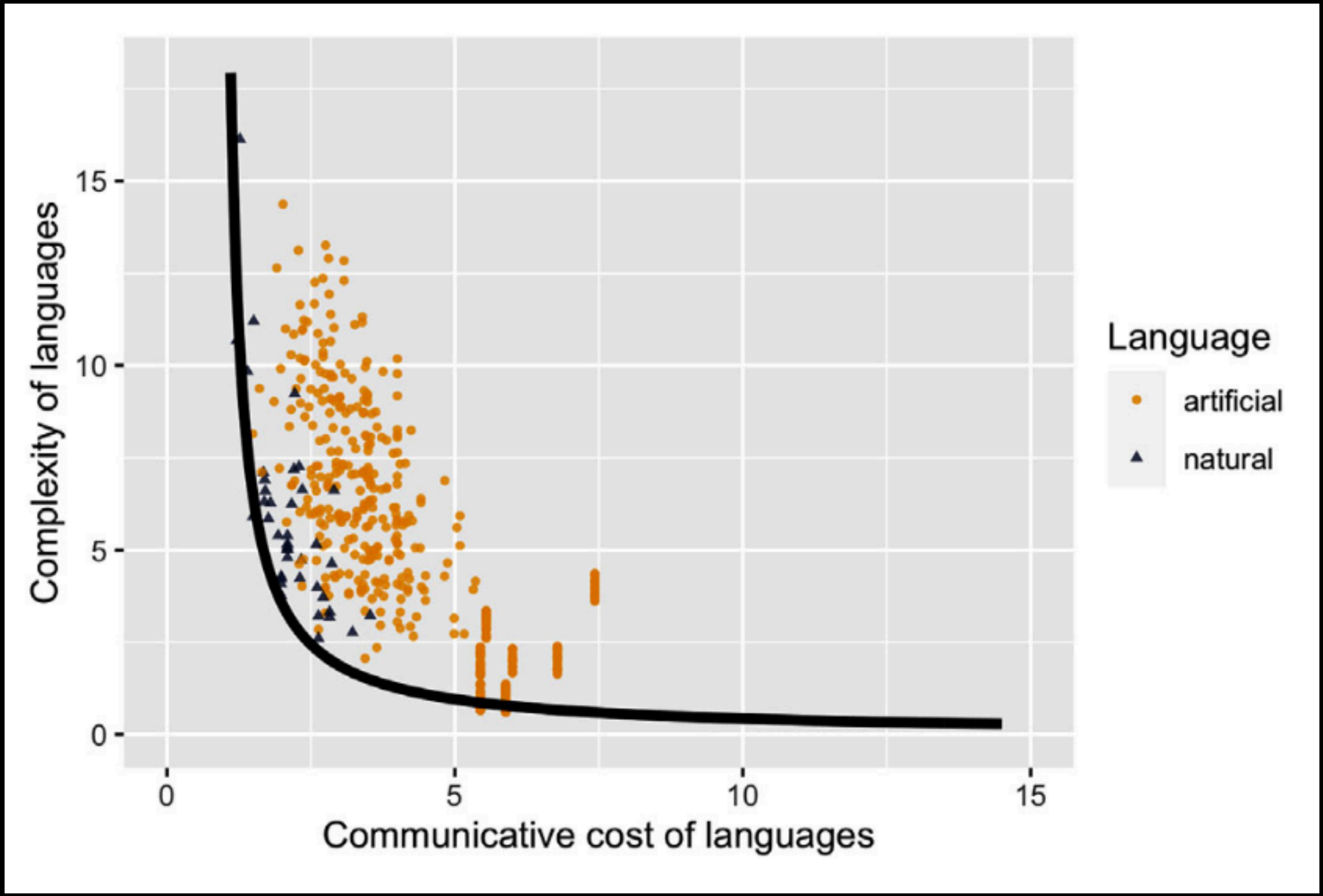
Complexity
1
0.75
0.50
0.25
0
0.7
0.8
0.9
1
Communicative Cost
Naturalness
1
0.75
0.50
0.25
0

Results: Degrees



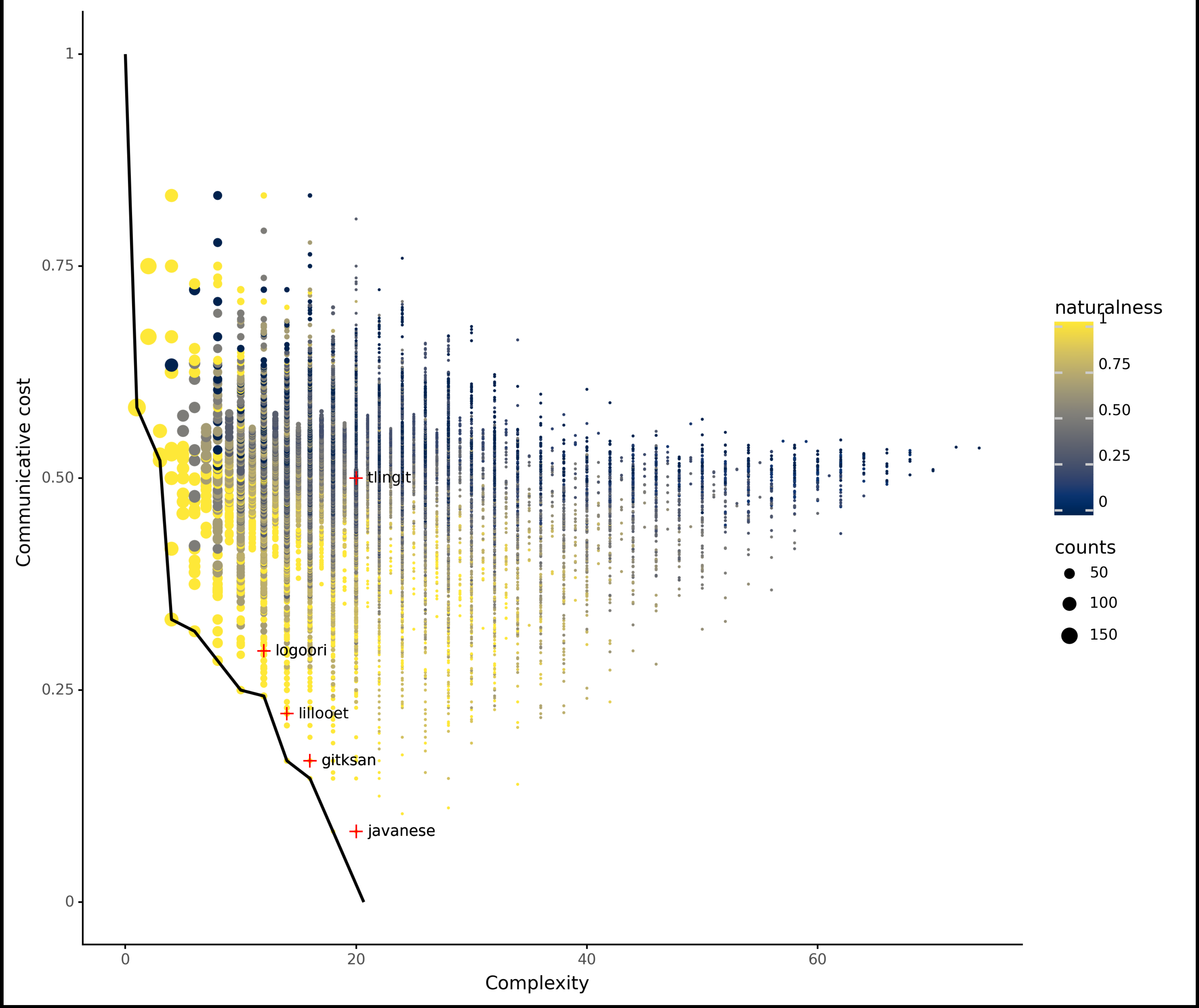
(a) Monotonicity (b) Conservativity

Indefinites



Denić, Steinert-Threlkeld, Szymanik 2022

Modals



Imel, Steinert-Threlkeld, 2022

Discussion Points

- Quantifiers:
  - Better sampling methods
  - Measuring learnability at scale as well
- Generally:
  - Learnability vs simplicity (cf van de Pol et al 2022)
  - Diachrony / language change (Carcassi et al ’22, Zaslavsky et al ’22)
  - Better cross-linguistic data (e.g. Keenan and Paper 2012/2017; Guo, Imel, S-T ’22)
  - General results about optimality and universals (e.g. Jäger 2017; Skinner, ongoing)
  - …

Journal of Semantics, 30, 2013: 315–334
doi:10.1093/jos/ffs014
Advance Access publication August 10, 2012

Conservativity and Learnability of Determiners

TIM HUNTER
Yale University

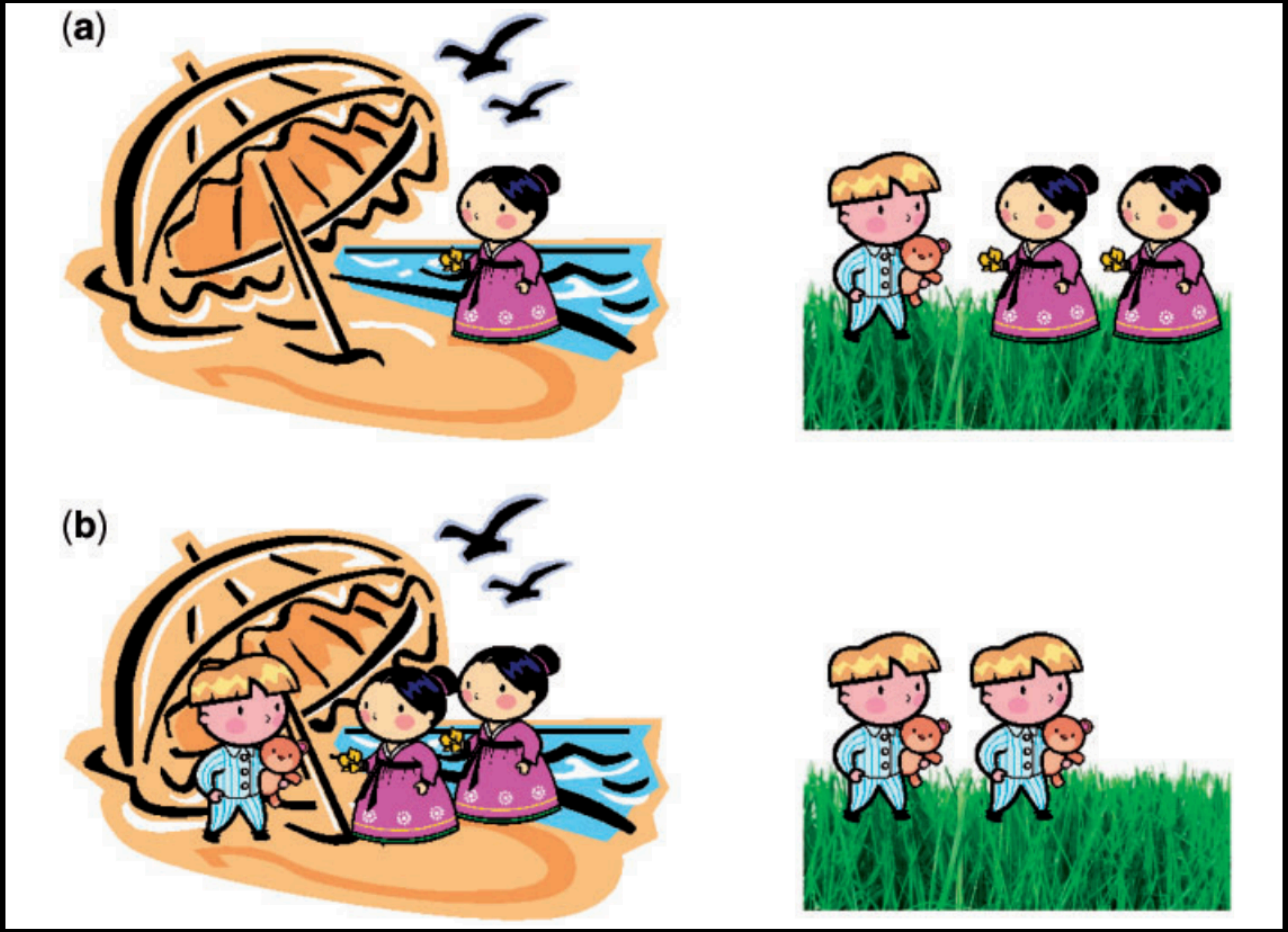
JEFFREY LIDZ
University of Maryland

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.

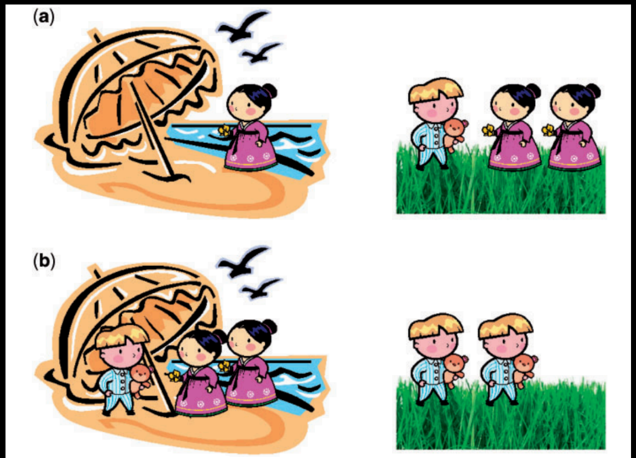
Gleeb vs Gleeb’

- Gleeb = ‘not all’
- Gleeb’ = ‘not only’



Hunter & Lidz, 2012

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)

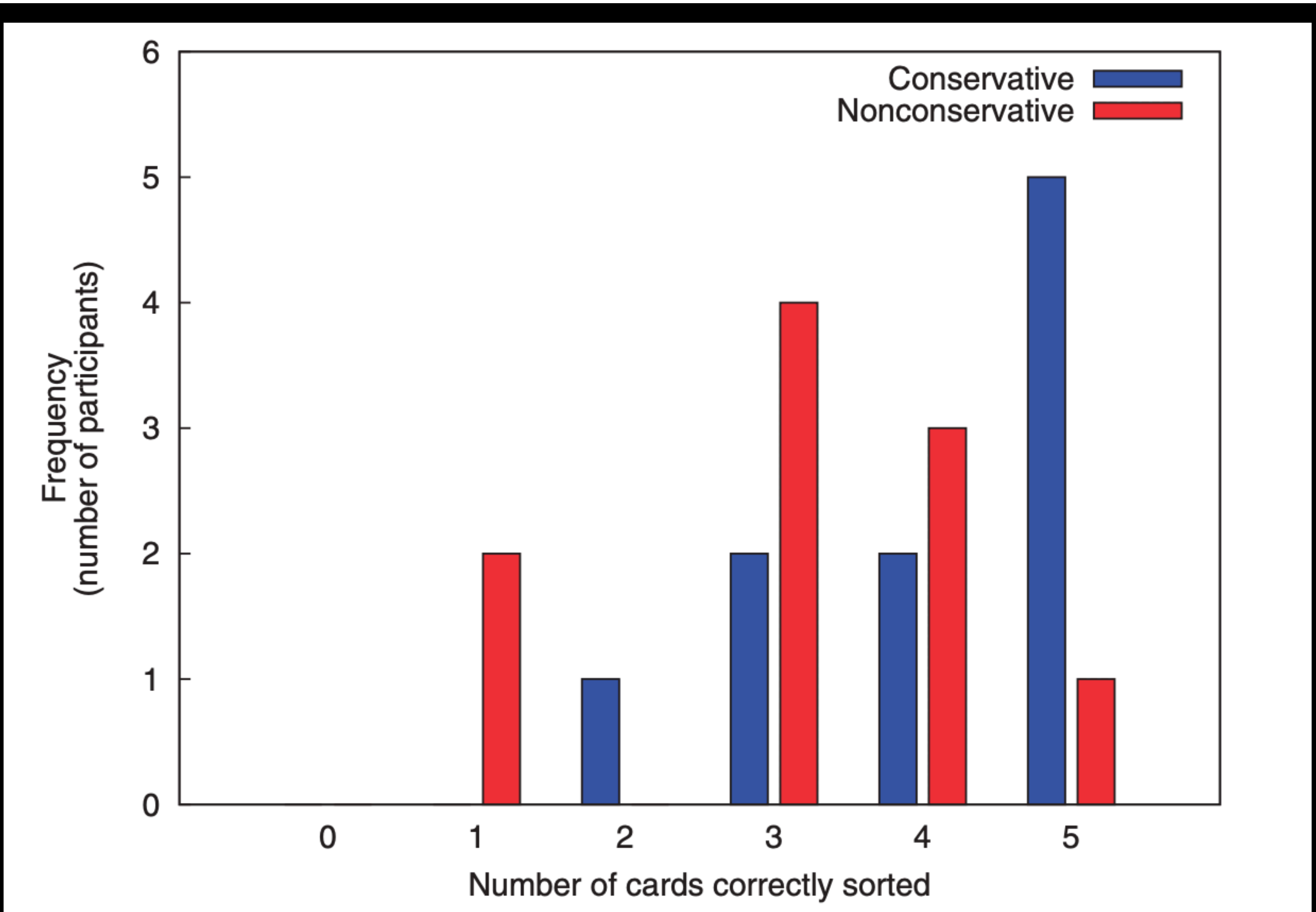
1. *The puppet told me that he likes this card because gleeb girls are on the beach*
2. *The puppet told me that he doesn't like this card because it not true that gleeb girls are on the beach.*

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%



Hunter & Lidz, 2012

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

–Perfect non-conservative kid; interpreting conservatively?

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? (Cf yesterday's minimal pair discussion)

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? (Cf yesterday's minimal pair discussion)
- Unclear that it replicates (Spenander and de Villiers 2019)

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? (Cf yesterday's minimal pair discussion)
- Unclear that it replicates (Spenander and de Villiers 2019)
- Also, what about the other universals?

Journal of Semantics, 00, 2019, 1–17
doi:10.1093/jos/ffz001
Advance Access Publication Date:
Article

OXFORD

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

Connectedness

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

Connectedness

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

Connectedness

- X is convex: $\forall a, b \in X, t \in [0,1]: \quad ta + (1-t)b \in X$
- *c* is between *a* and *b*: $\exists t \in [0,1]: \quad c = ta + (1-t)b$
- 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: $\forall a, b \in X, t \in [0,1]: \quad ta + (1-t)b \in X$
- *c* is between *a* and *b*: $\exists t \in [0,1]: \quad c = ta + (1-t)b$
- 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 <et, t>) 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 <et, t>) is connected:
  if $A \subset C \subset B$ then $Q(C) \geq Q(A)$ or $Q(C) \geq Q(A)$
- 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 <et, t>) is connected:
  if $A \subset C \subset B$ then $Q(C) \geq Q(A)$ or $Q(C) \geq Q(A)$

- 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.

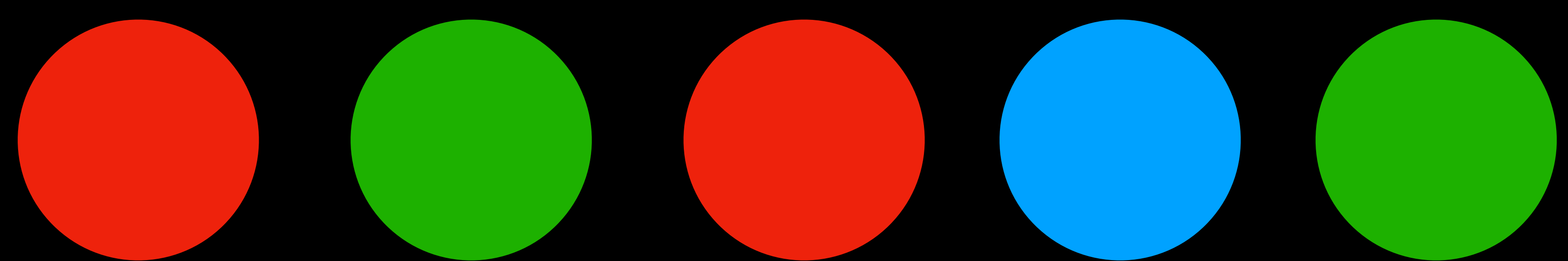
All word meanings are connected. (?)

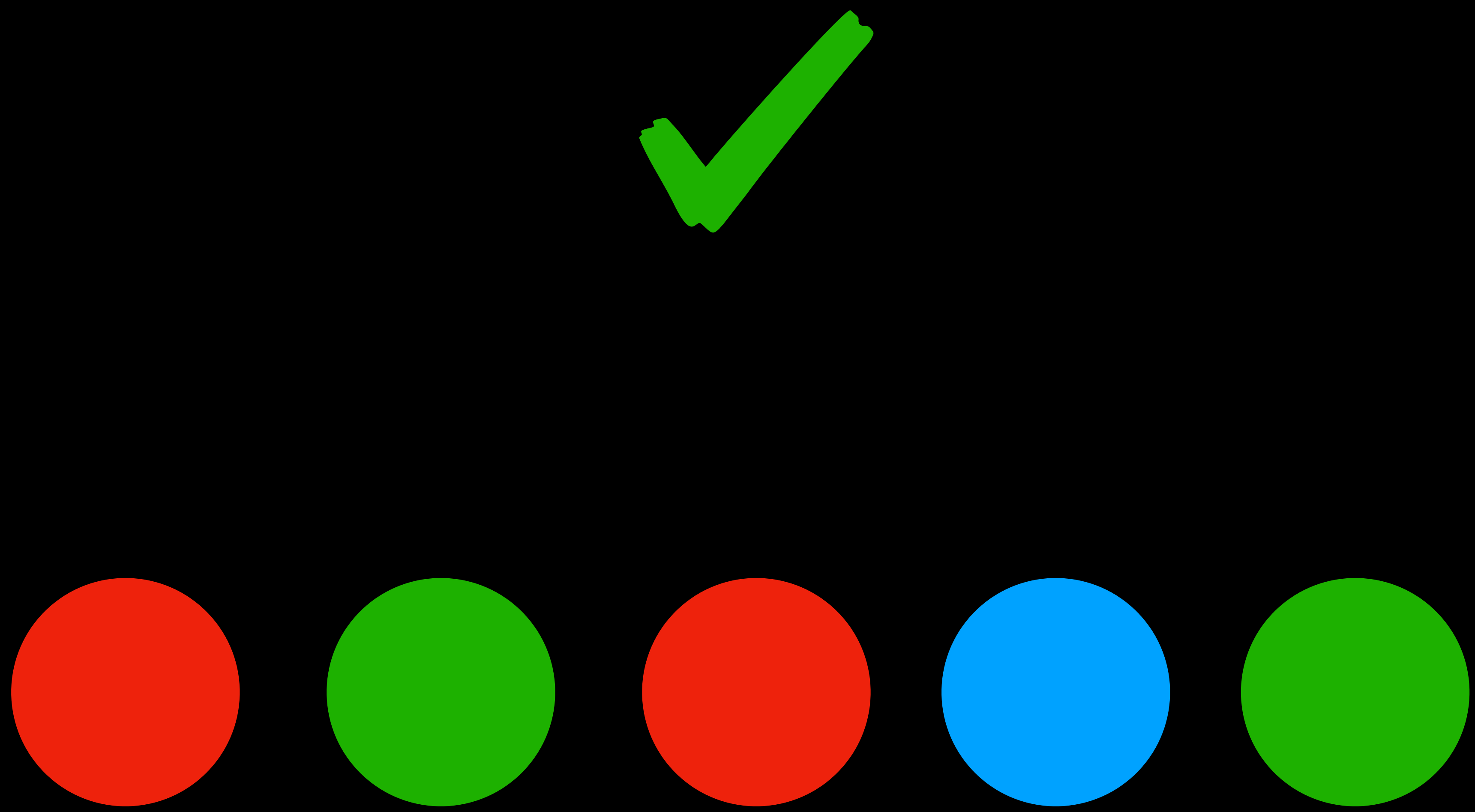
“Connected quantifiers are typically expressed in simpler ways than non-connected quantifiers.”

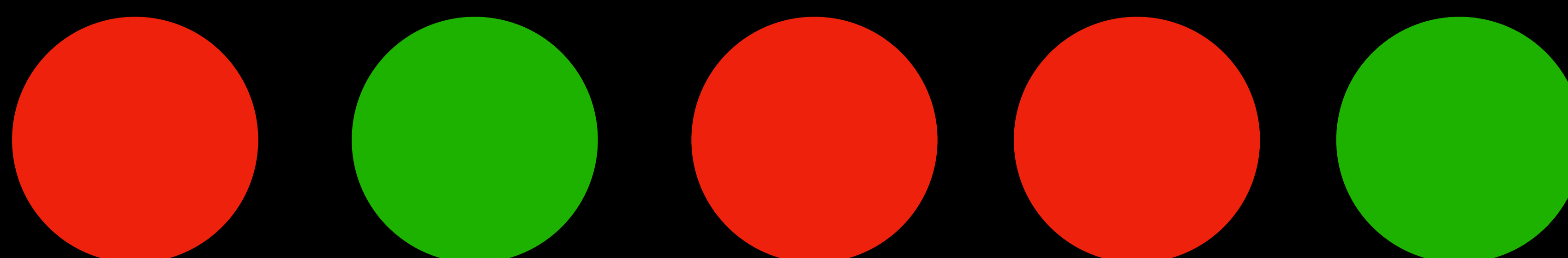
"Connected quantifiers are typically expressed in simpler ways than non-connected quantifiers."

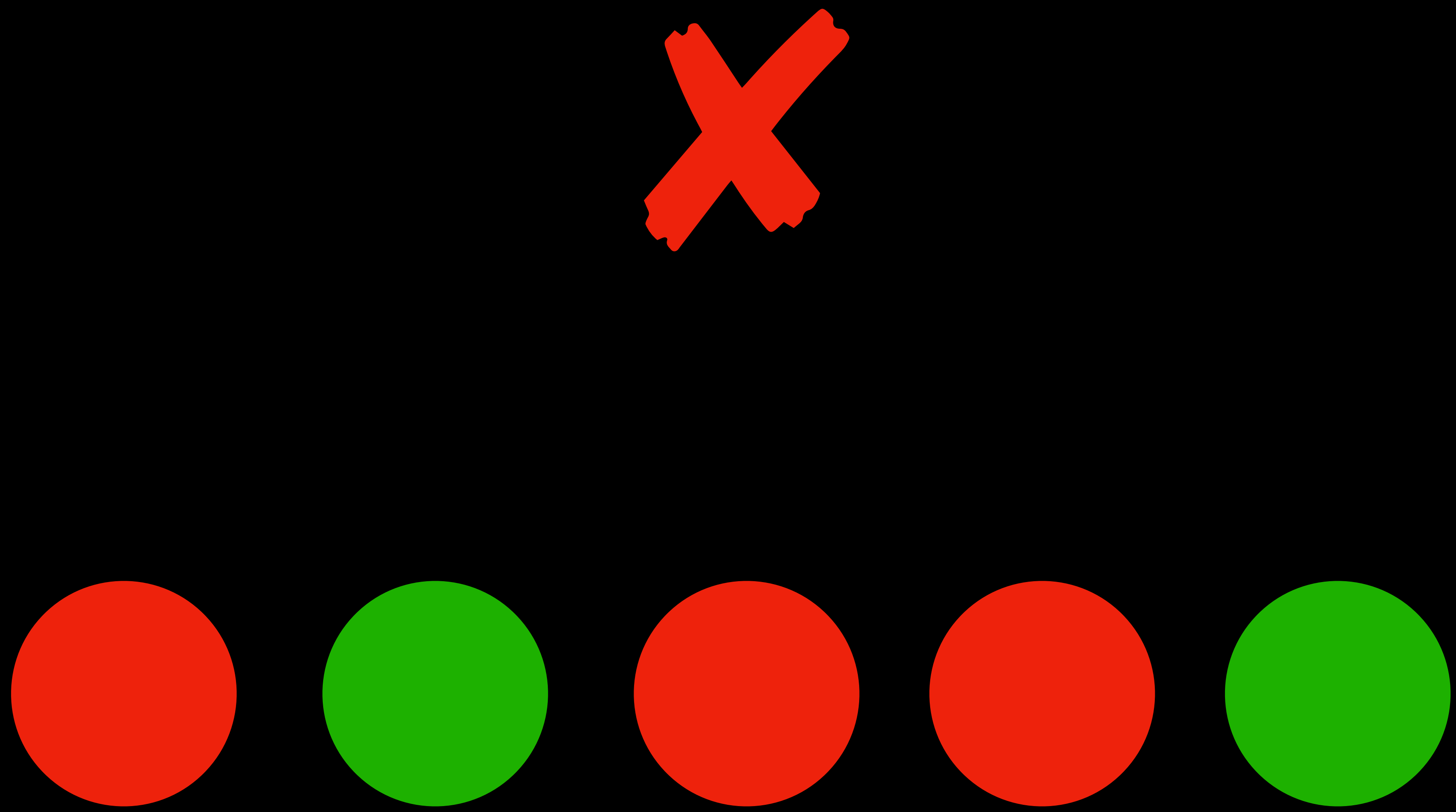
Compare, e.g., 'between 5 and 10' or '5 to 10' with 'less than 5 or more than 10'.

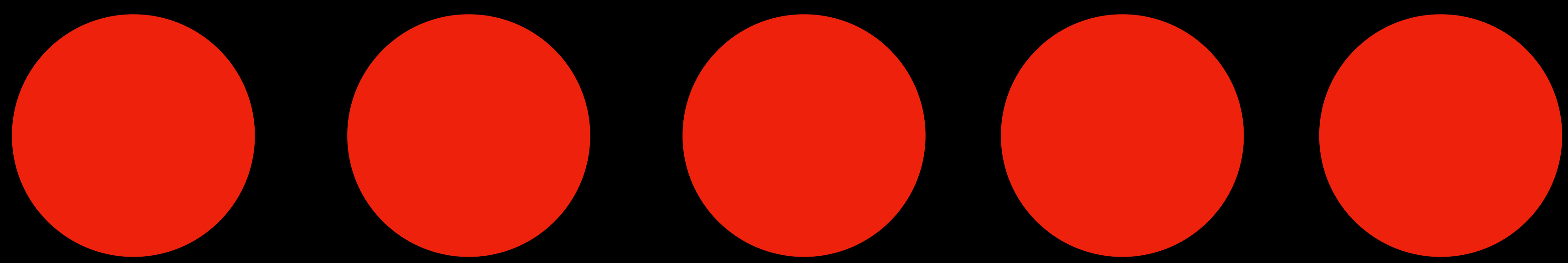
Experiment Time!

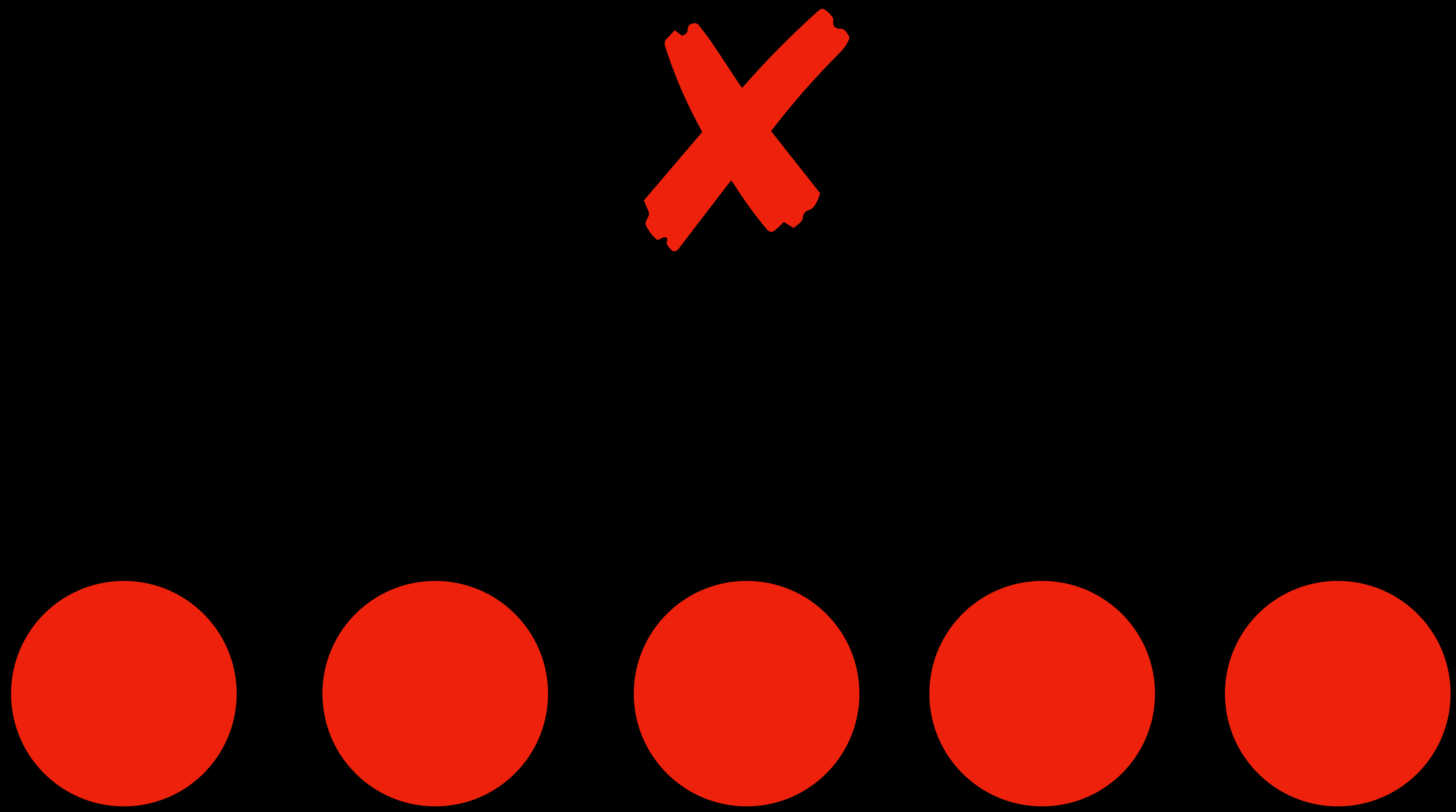


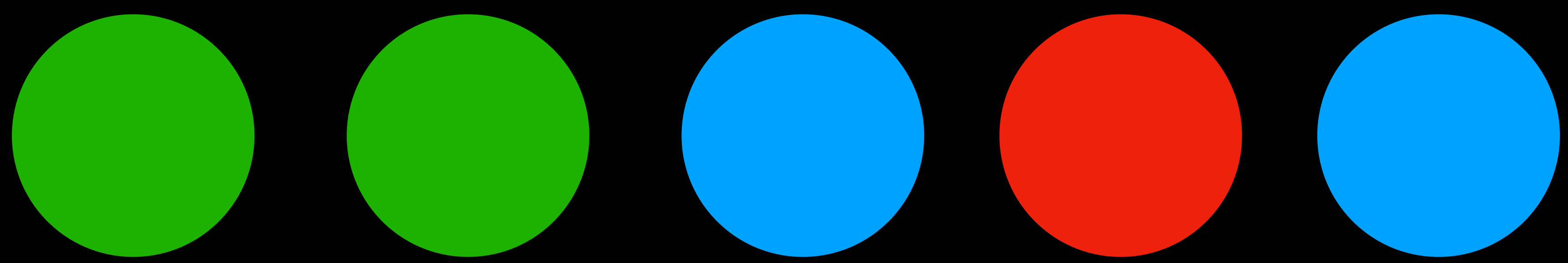


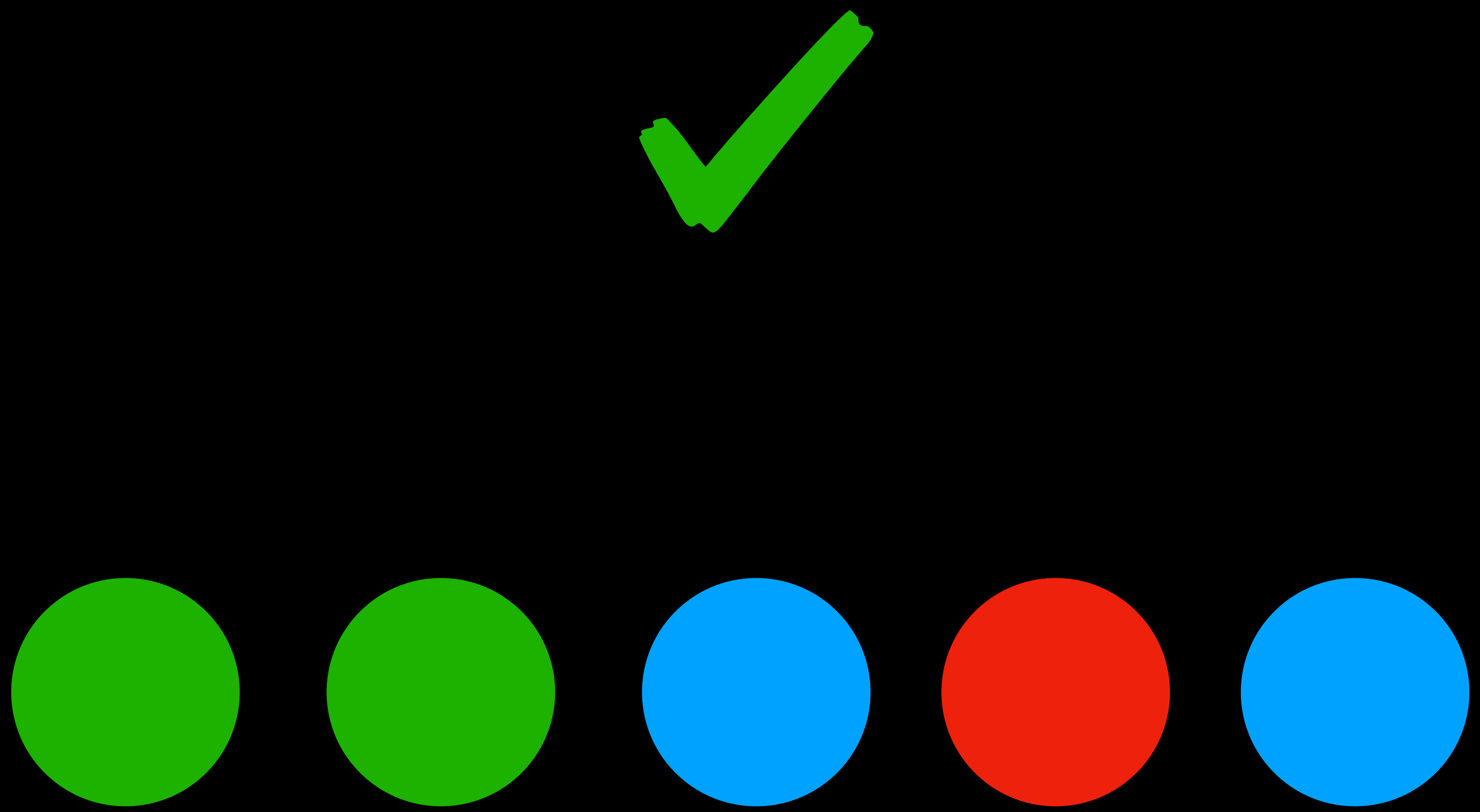


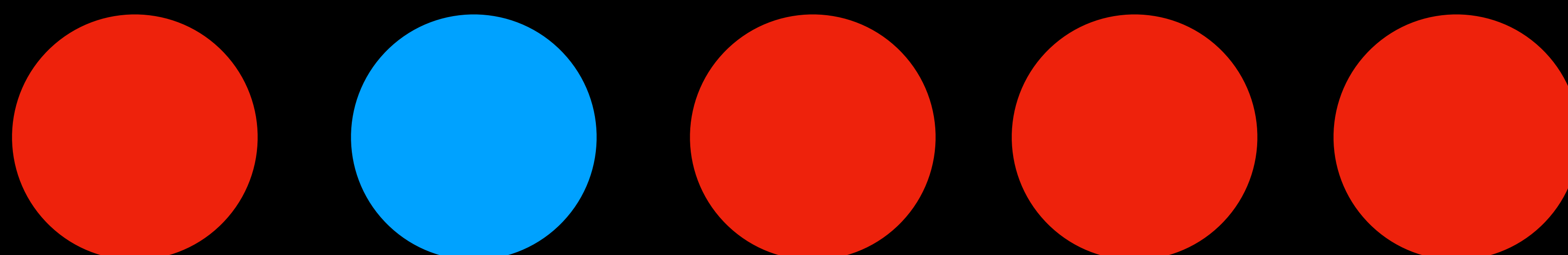


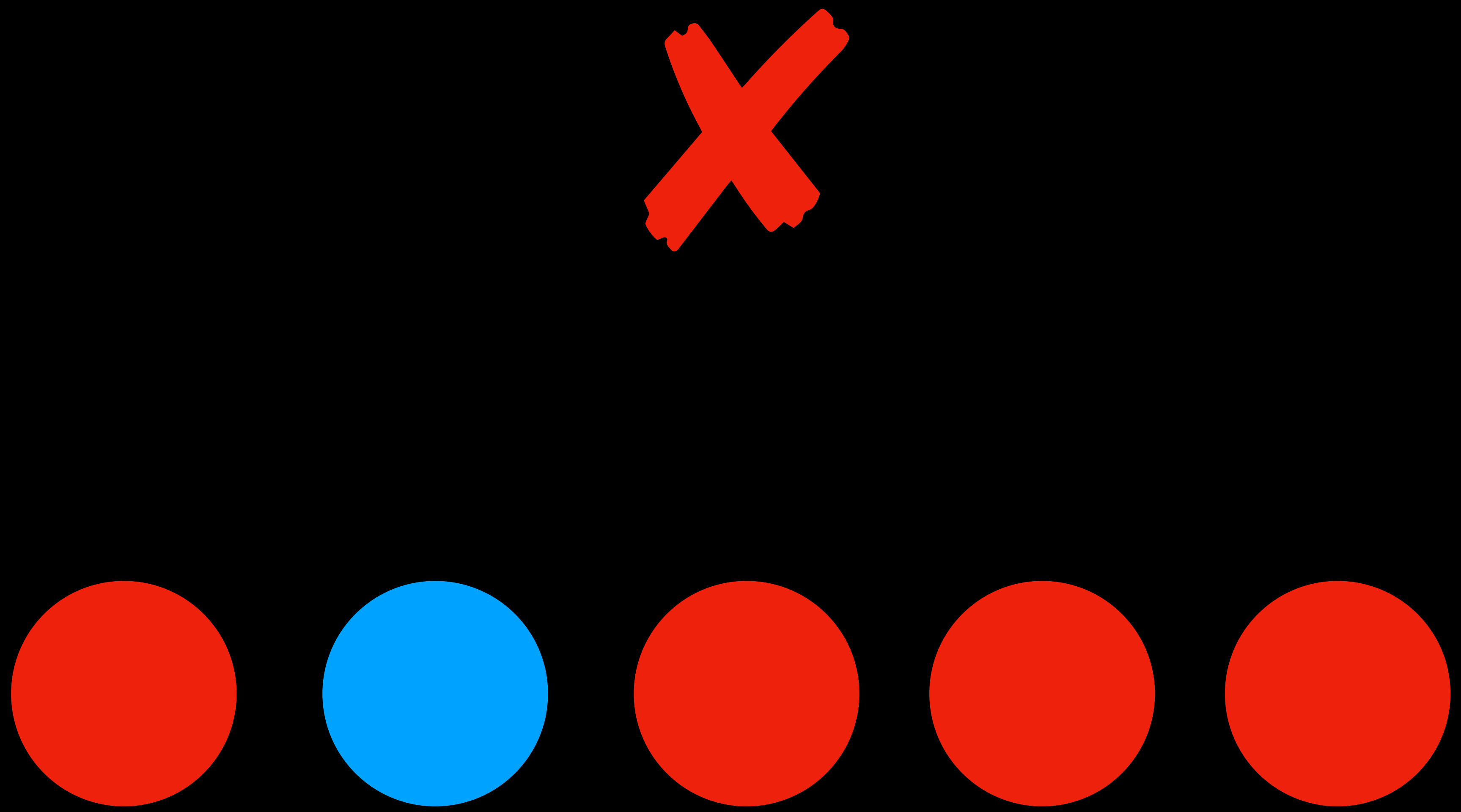


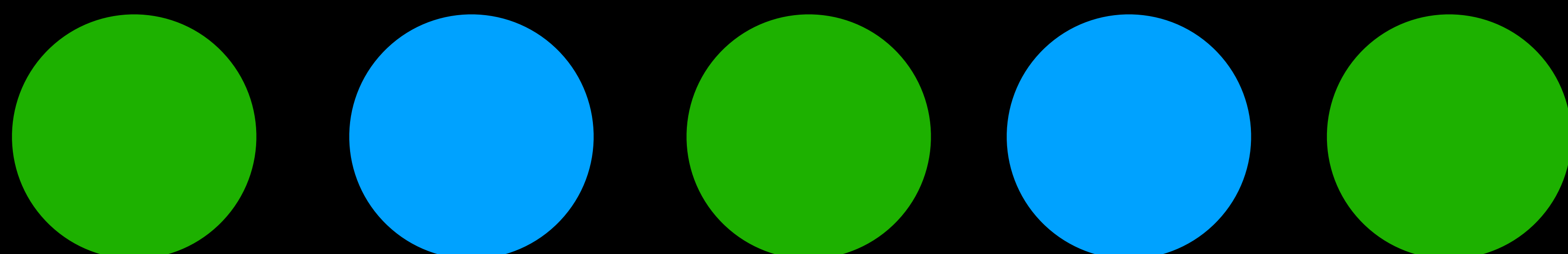


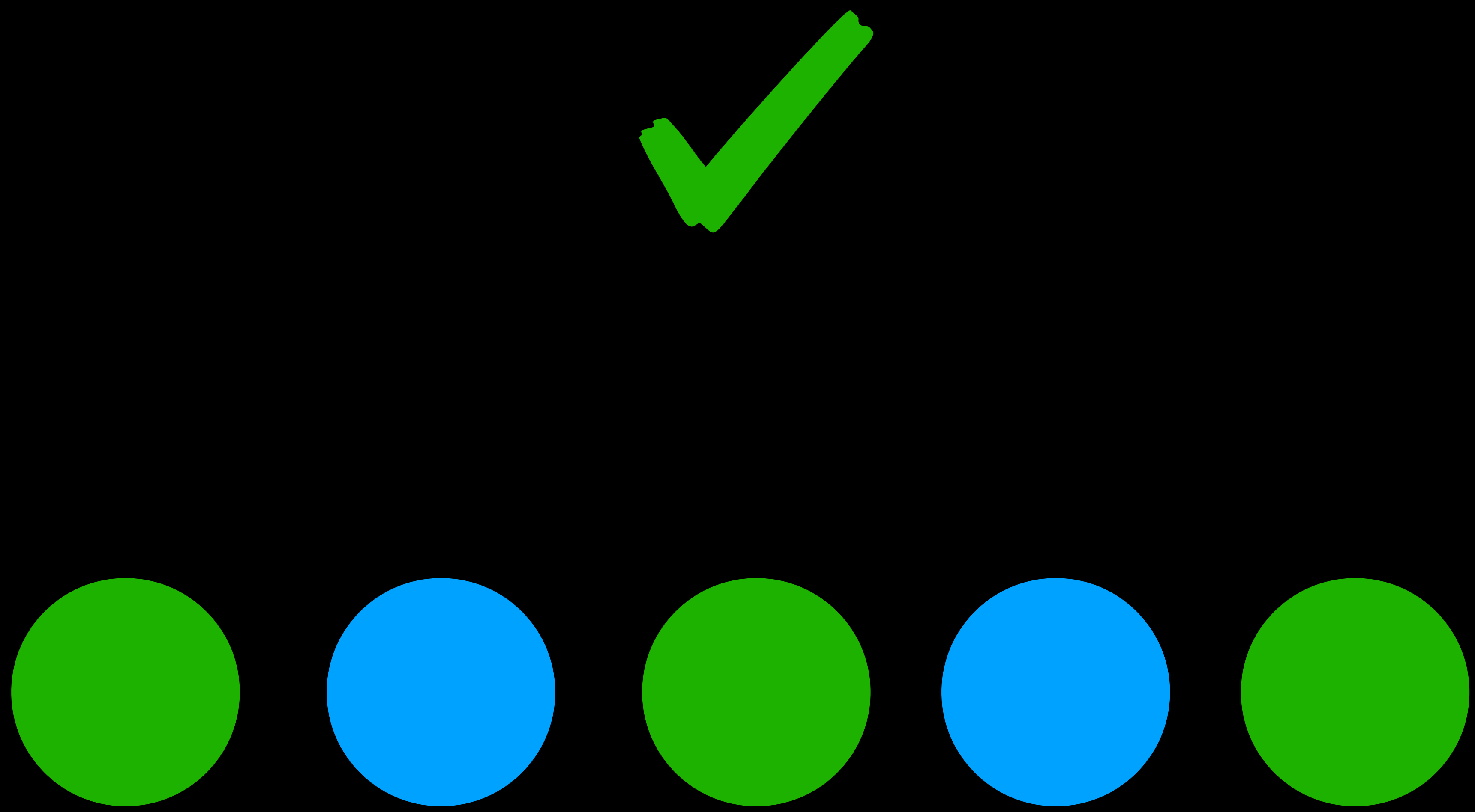




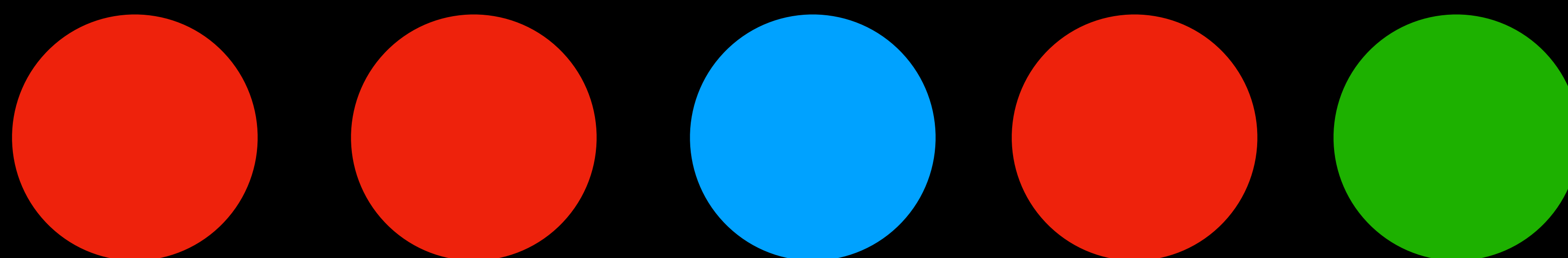


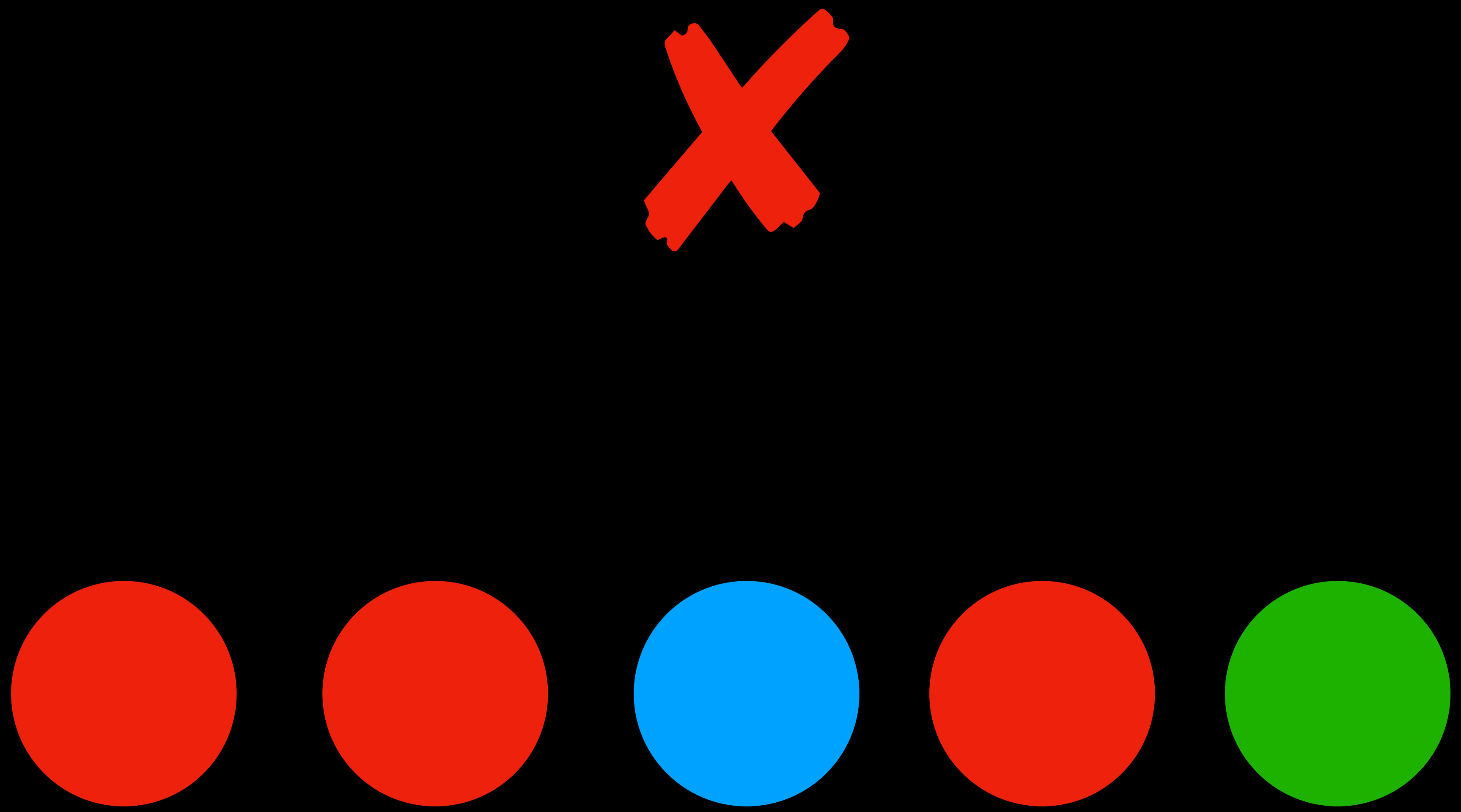


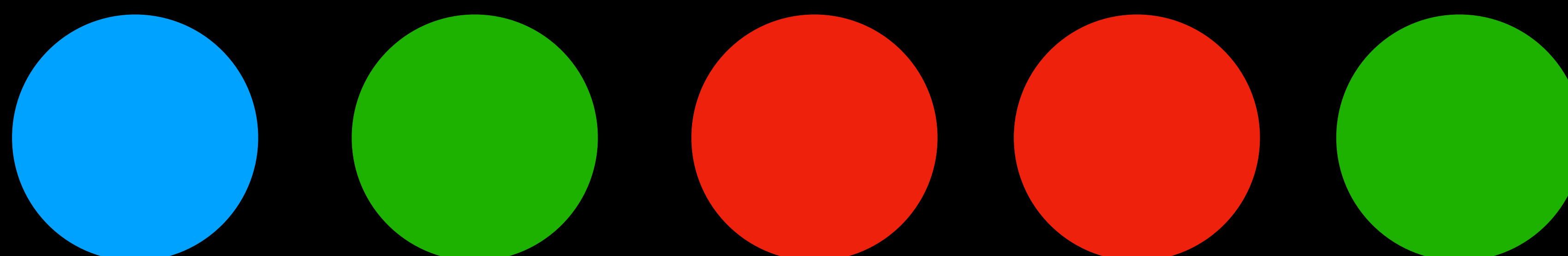


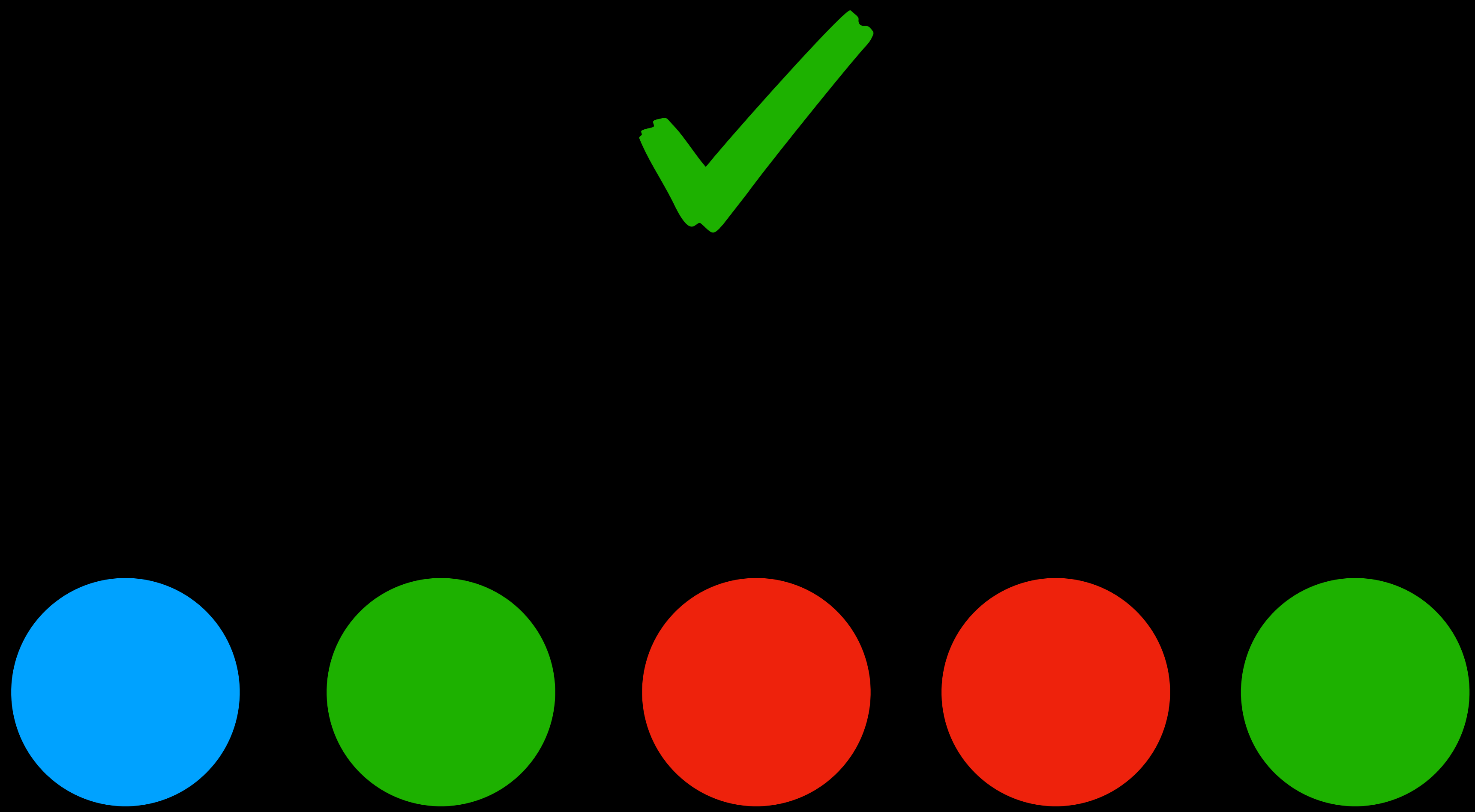


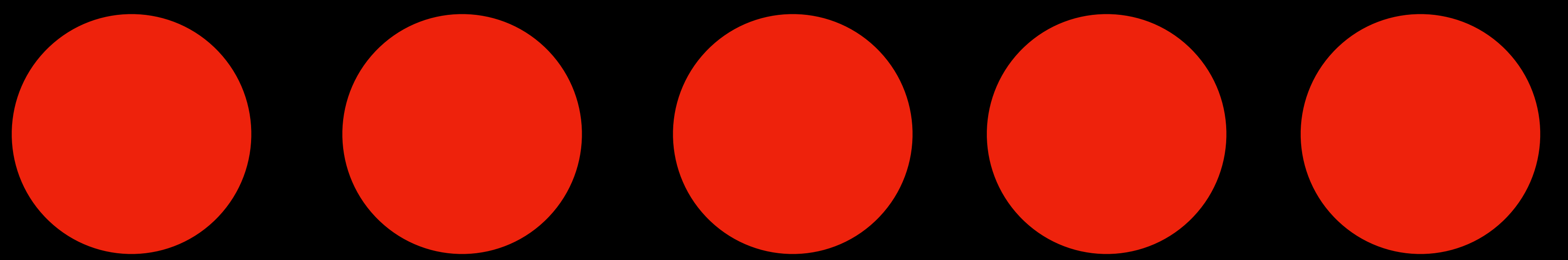
At most 2 red dots.

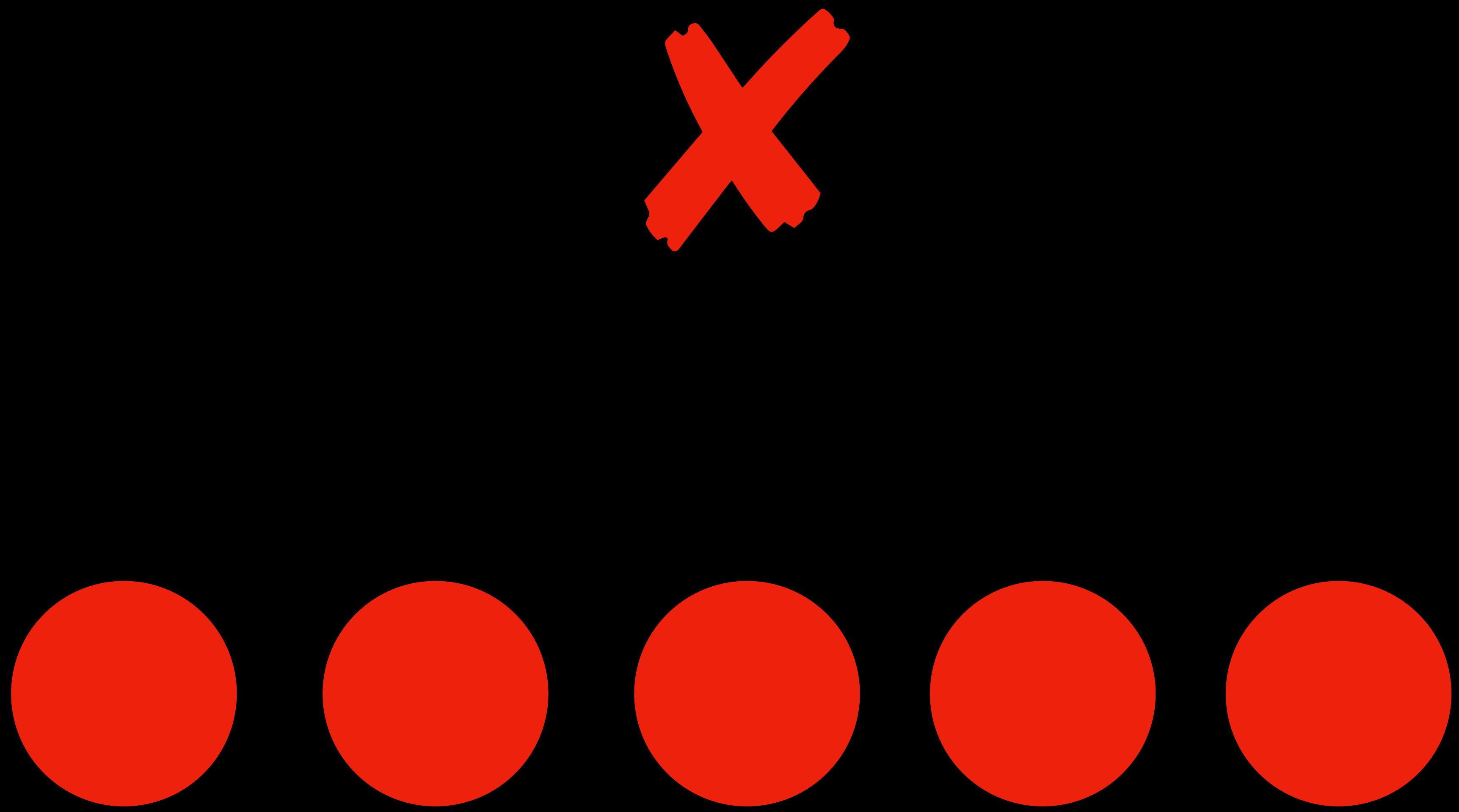


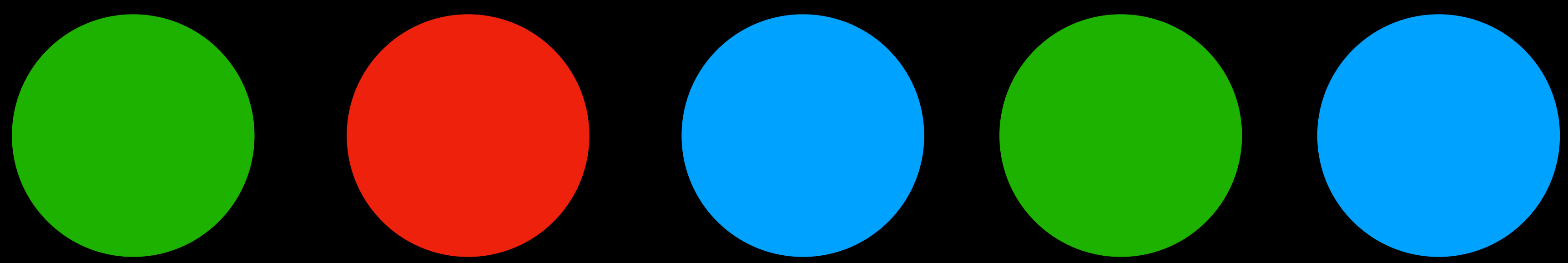


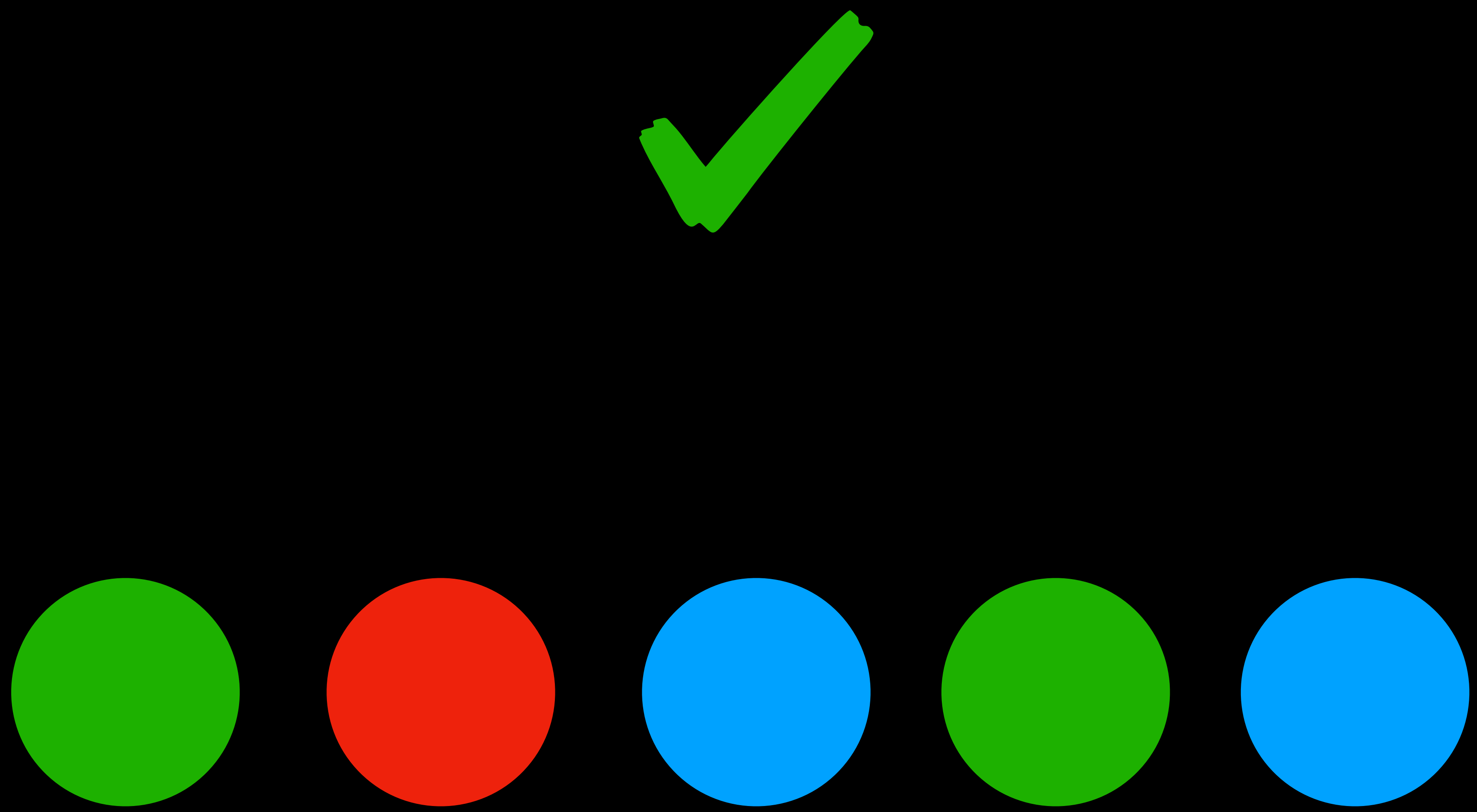


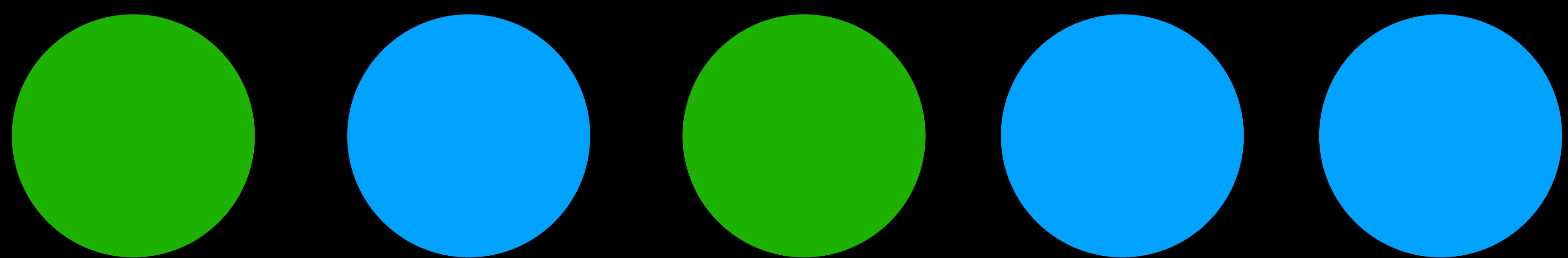


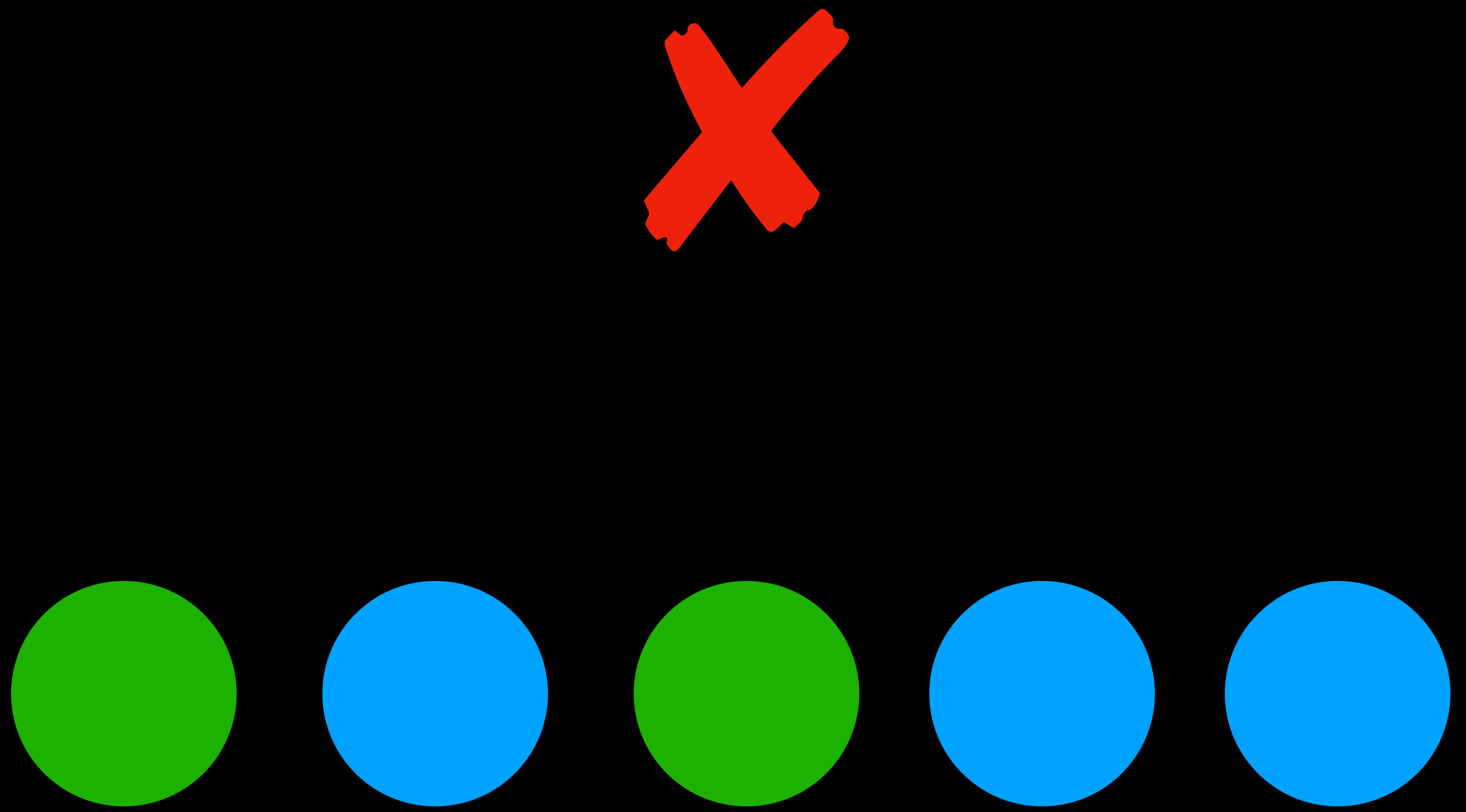


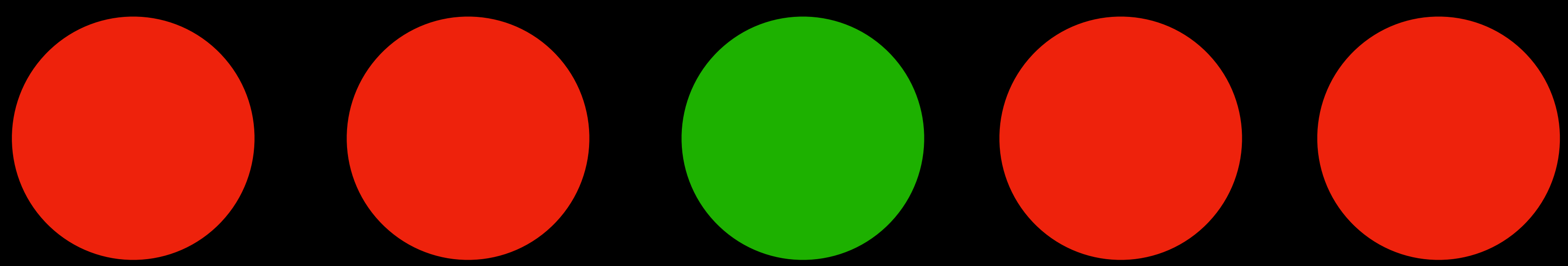


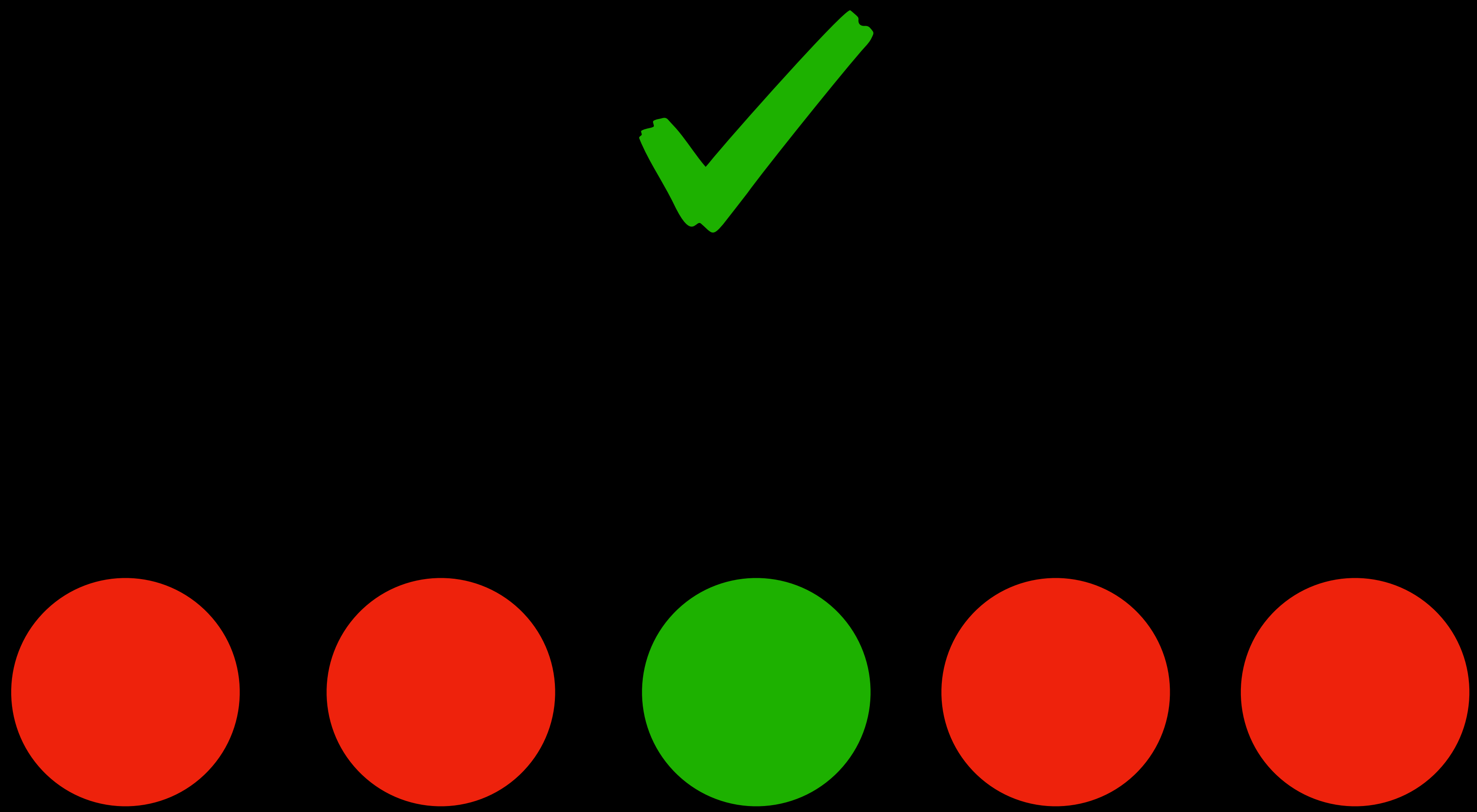










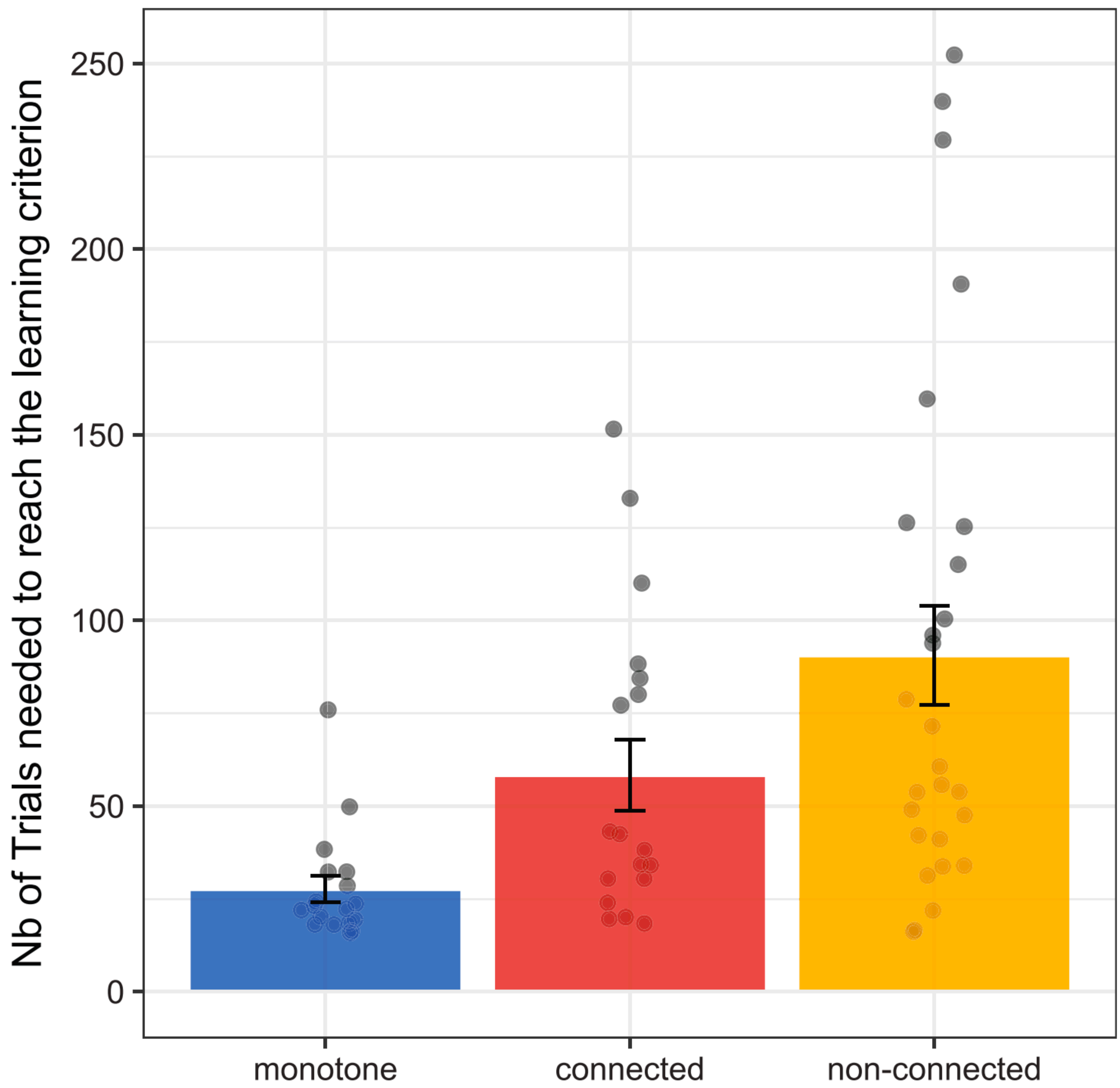


1, 2, or 4 red dots.

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
250
200
150
100
50
0
monotone
connected
non-connected

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]

[a]Laboratoire 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; [b]Centre for Language Evolution, University of Edinburgh, Edinburgh EH8 9YL, United Kingdom; [c]Laboratory of Cognitive Psychology, Aix-Marseille University, CNRS, 13331 Marseille Cedex 3, France; and [d]Institute 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

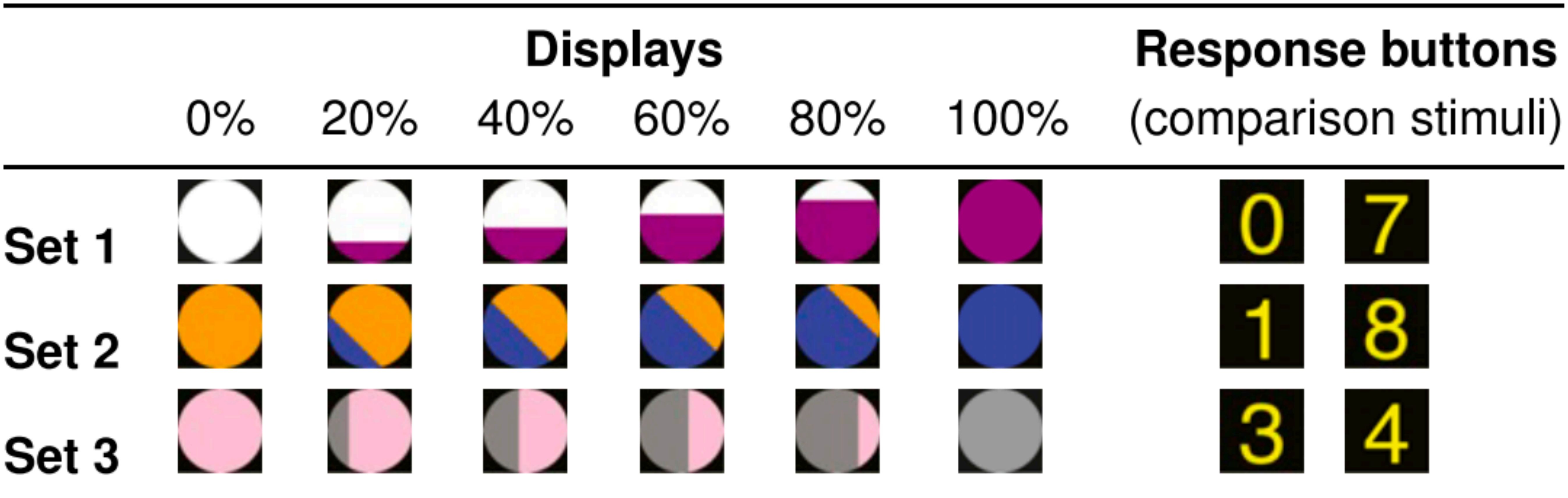
Humans and animals categorize objects in the world into

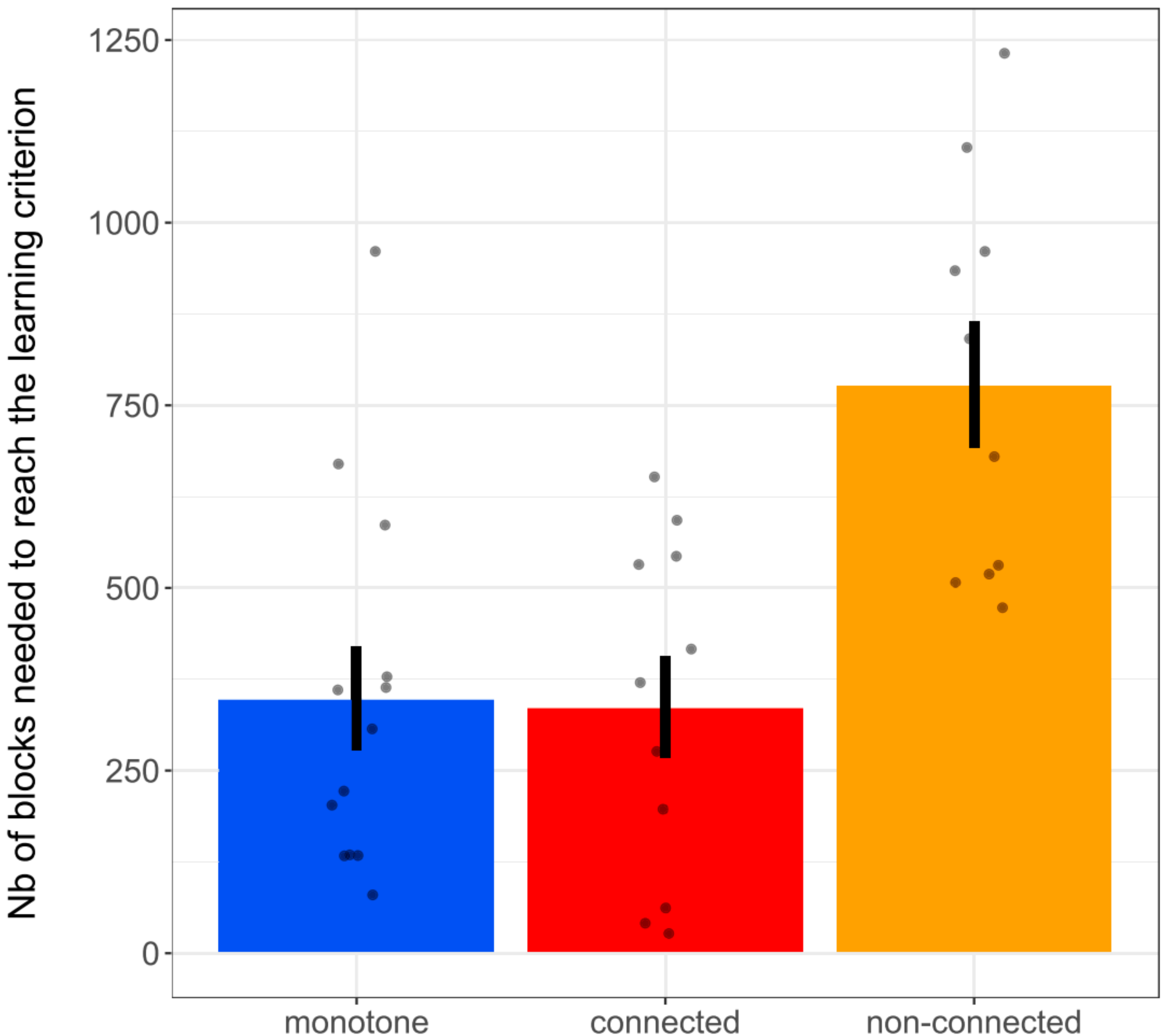
show that humans have corresponding learning biases favoring connected quantifiers, as evidenced by performance on rule learning, or pattern extraction, tasks: It is easier to discover connected rules than nonconnected ones, and easier still to discover monotone ones.

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

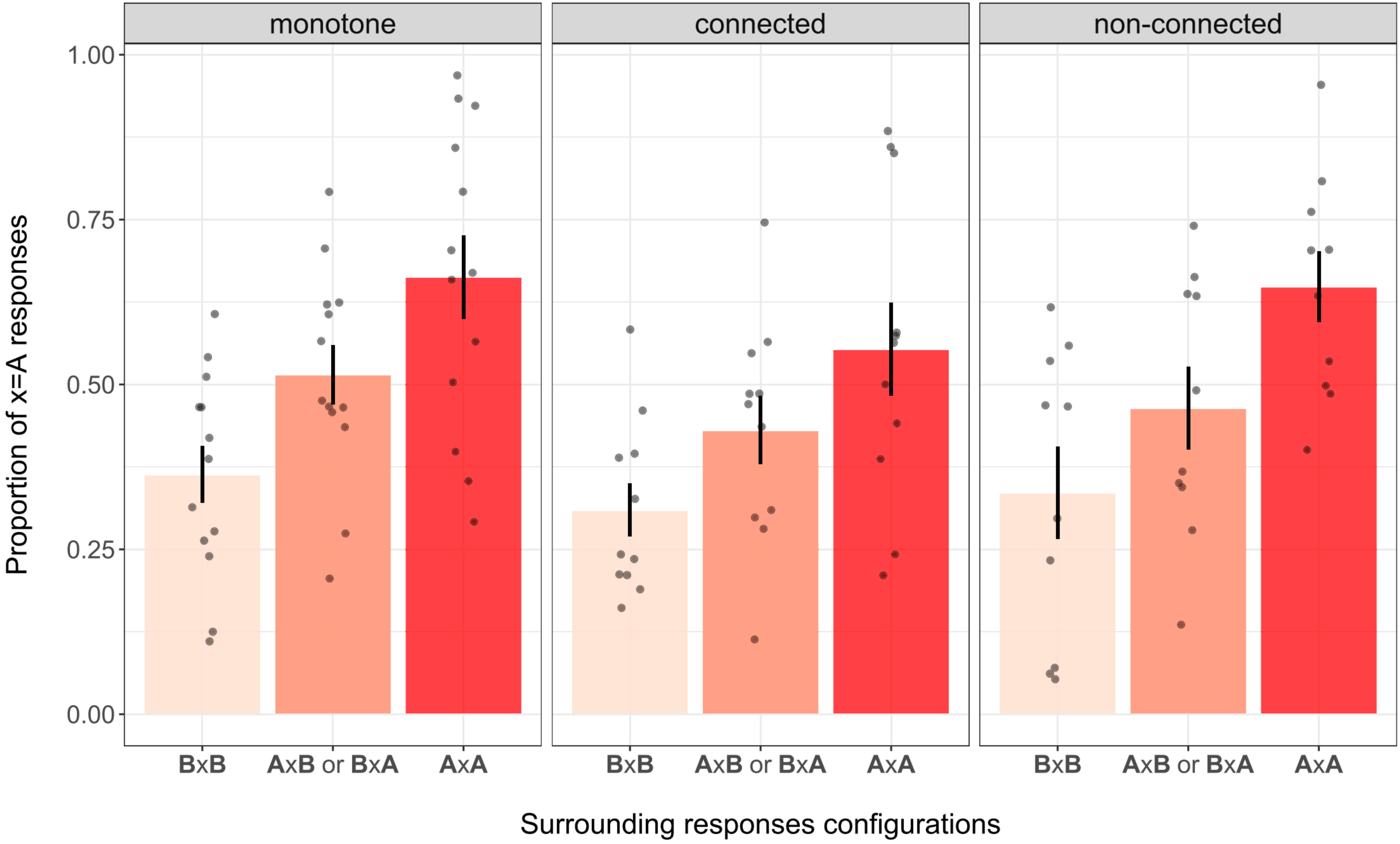
Significance

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
0%
20%
40%
60%
80%
100%
Response buttons
(comparison stimuli)
Set 1
Set 2
Set 3
0
7
1
8
3
4



Nb of blocks needed to reach the learning criterion
0
250
500
750
1000
1250
monotone
connected
non-connected



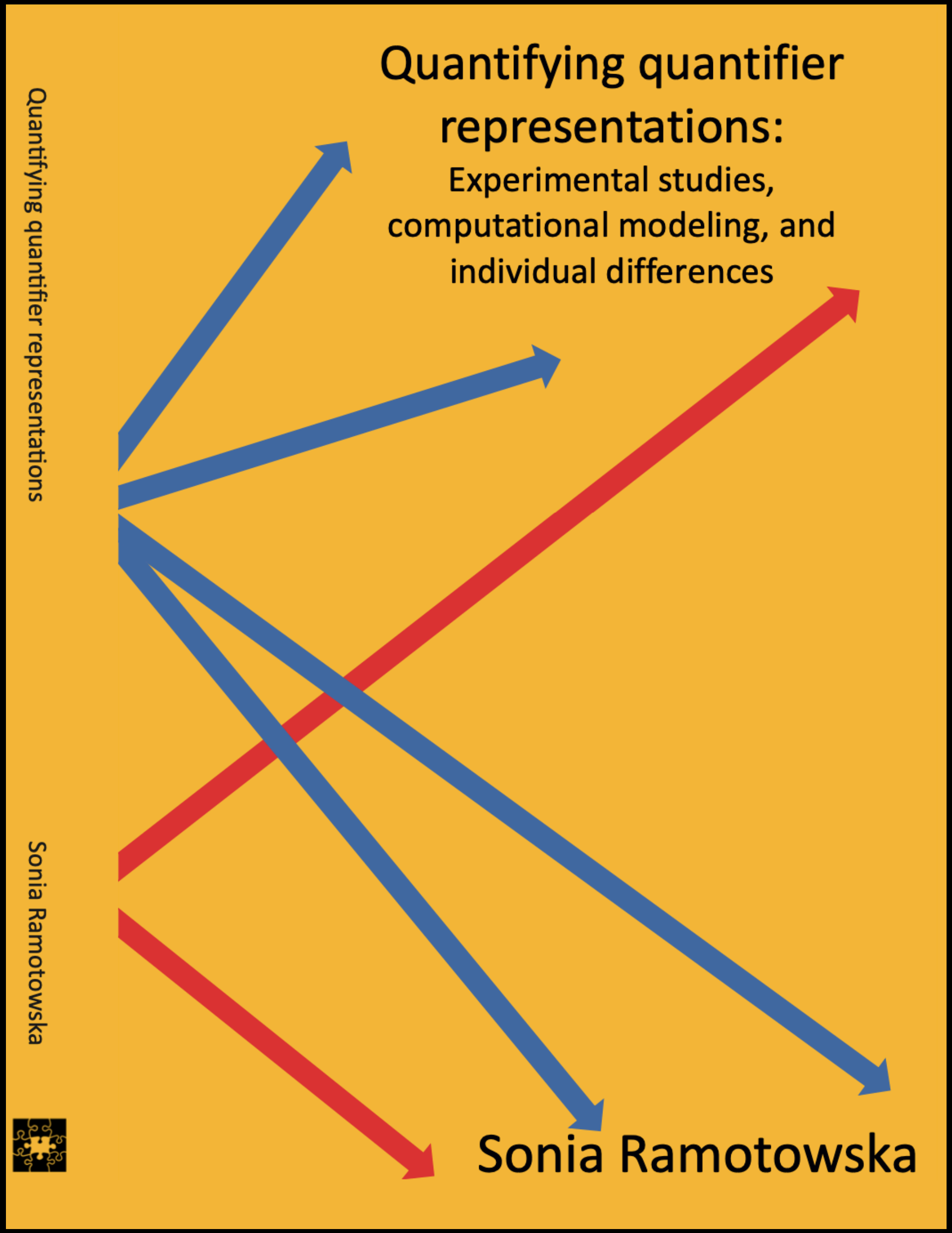
monotone
connected
non-connected
1.00
0.75
0.50
0.25
0.00
Proportion of x=A responses
BxB
AxB or BxA
AxA
Surrounding responses configurations

“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?

Large Scale Learnability Experiment

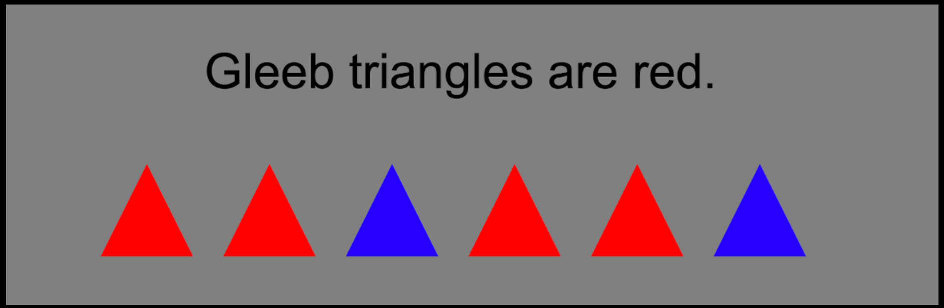


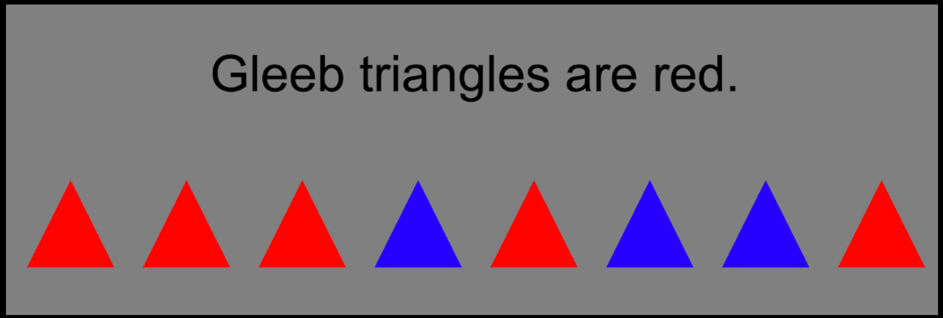
Tested Quantifiers

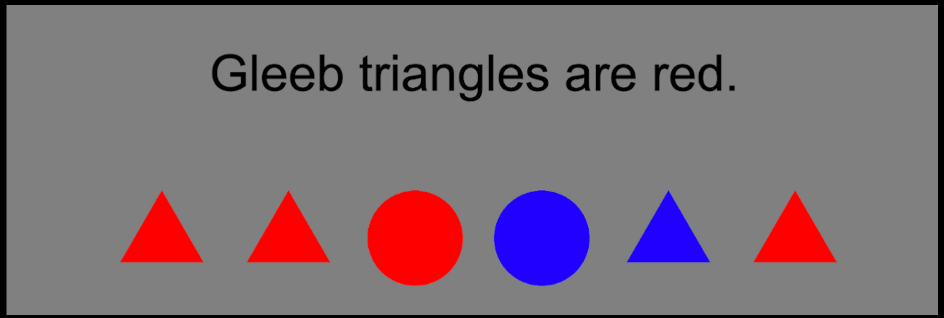
- at least 3 & at most 2 vs. between 3 and 6 & at most 2 or at least 7
- between 3 and 6 vs. at most 2 or at least 7
- at least 3 & at most 3 vs. first and the last 3
- not all vs. not only

Design

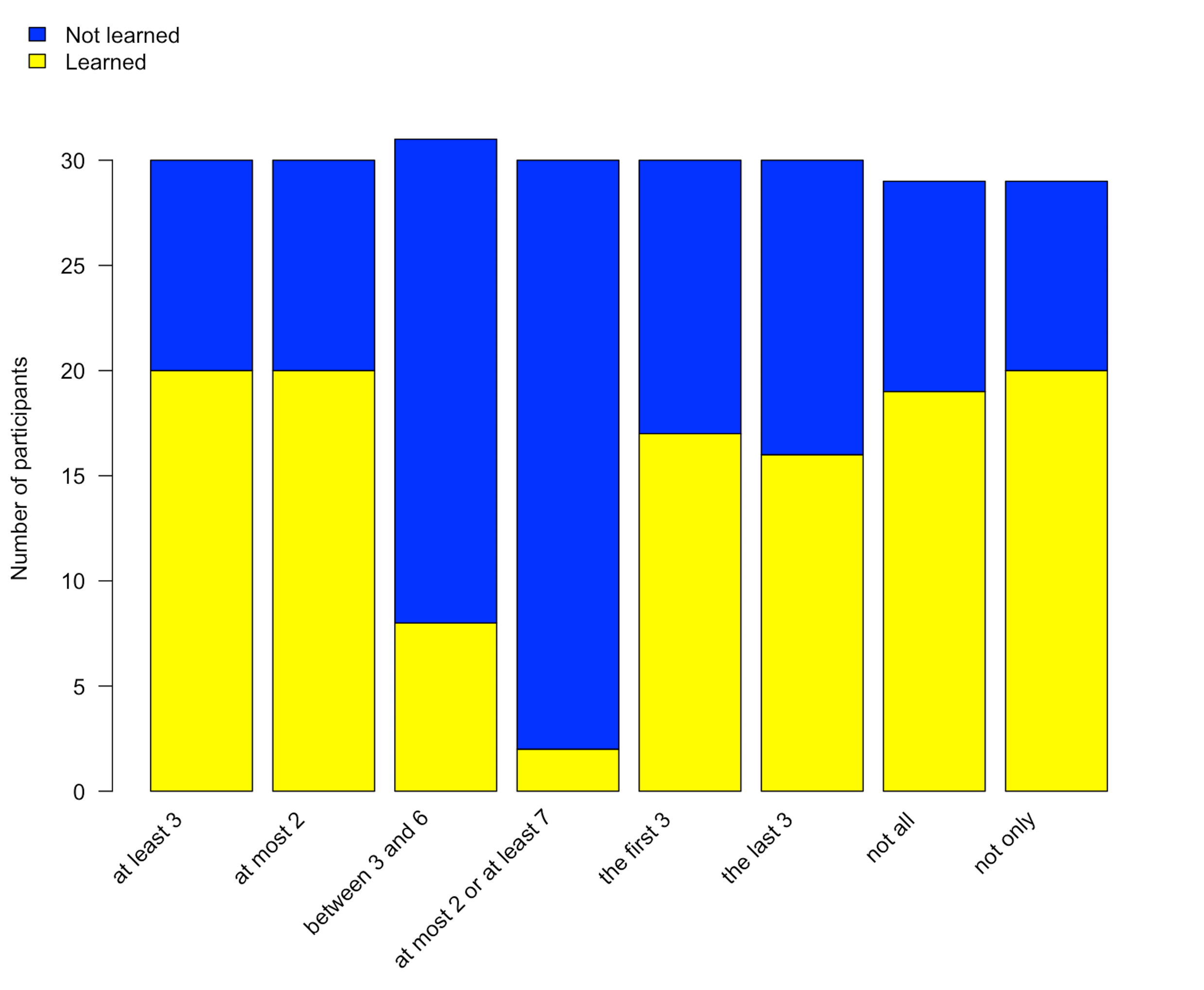
- ~30 participants for each quantifier (H&L:10, S&dV:9)
- 96 trials, 8 implicit blocks for 12 trials.





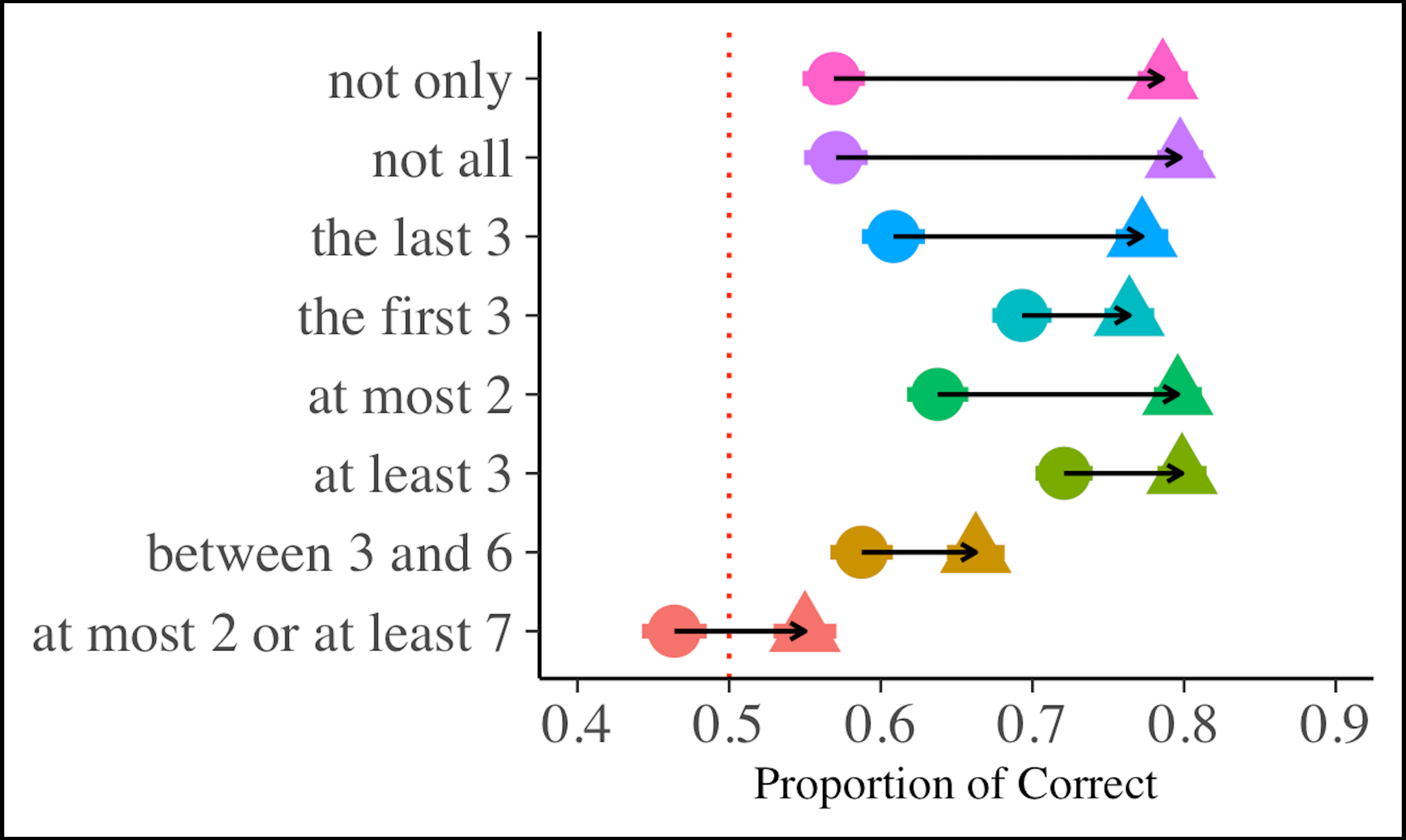


#participants performing above chance in the last block



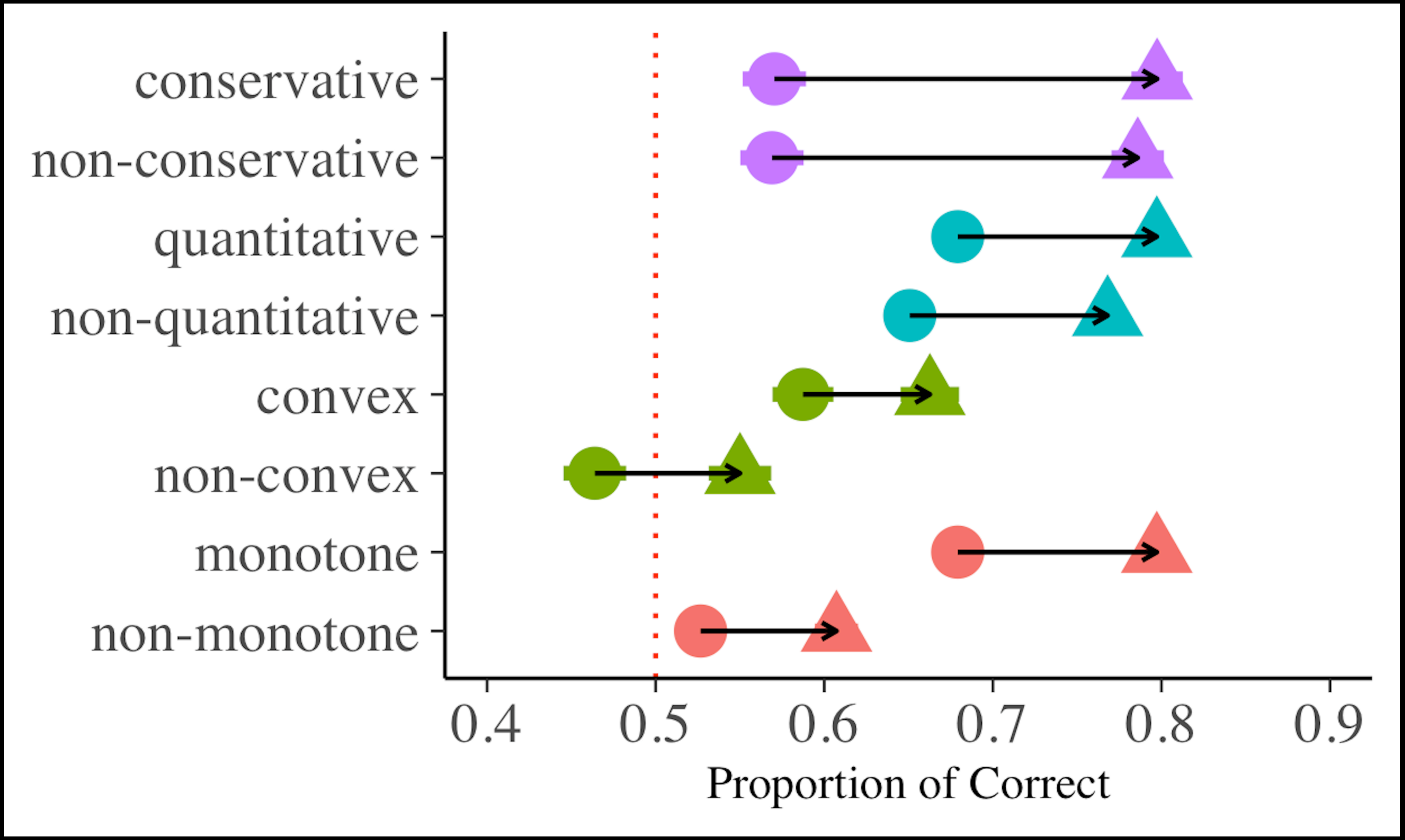
Ramotowska, van Maanen, Szymanik, 2022

Mean accuracies in the first 25% and the last 25% of the trials for each quantifier



Ramotowska, van Leendert, Szymanik, 2022

Mean accuracies in the first 25% and the last 25% of the trials for each universal



Ramotowska, van Maanen, Szymanik, 2022

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?

LANGUAGE AND COGNITIVE PROCESSES
2010, 25 (1), 130–148

Psychology Press
Taylor & Francis Group

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.

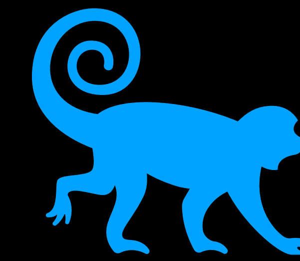
Superlative vs. comparative quantifiers

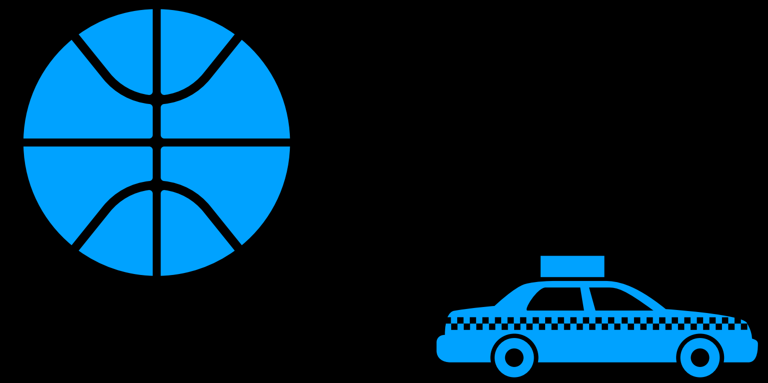
- “at least 3” = “more than 2”?
- “at most 3” = “fewer than 4”?
- Geurts and Nouwen (2007): No, superlatives are more complex

1. **Berta didn’t have at most 3 martinis.*
2. *Berta didn’t have fewer than 4 martinis.*

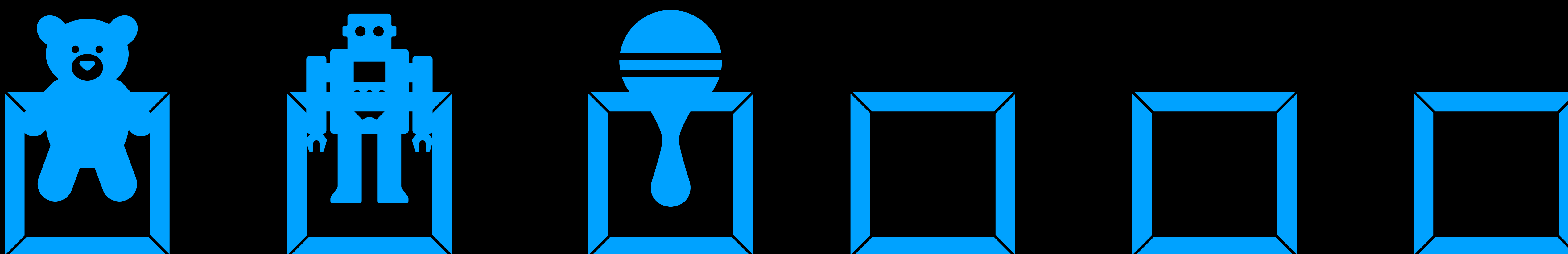
“At most n A are B” means that the speaker – considers it possible that there is a set of n A’s that are B, and – is certain that there is no larger set of A’s that are B.

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 response	%	3-arrangement response	%	4-arrangement response	%	% all correct (sd)
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	
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	

Geurts et al. 2015

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 Hrzica[e], Kleanthes K. Grohmann[f], Athina Skordi[g], Kristine Jensen de López[h], Lone Sundahl[h], Angeliek van Hout[i], Bart Hollebrandse[i], Jessica Overweg[i], Myrthe Faber[i,j], Margreet van Koert[k], Nafsika Smith[l], Maigi Vija[m], Sirli Zupping[m], Sari Kunnari[n], 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 Kapalková[ii], 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]

[a]Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge CB3 9DA, United Kingdom; [b]Department of Linguistics and English Language, University of Edinburgh, Edinburgh EH8 9AD, United Kingdom; [c]Department of Linguistics and Basque Studies, University of the Basque Country, ES-01006 Vitoria-Gasteiz, Spain; [d]Department of Catalan Philology, Universitat Autònoma de Barcelona, 08913 Bellaterra, Spain; [e]Department of Speech and Language Pathology, University of Zagreb, 10000 Zagreb, Croatia;[f]Department of English Studies, University of Cyprus, 1678 Nicosia, Cyprus; [g]Independent researcher; [h]Center for Developmental and Applied Psychological Science, Aalborg University, DK 9220 Aalborg East, Denmark; [i]Center for Language and Cognition Groningen, University of Groningen, 9700 AS Groningen, The Netherlands; [j]Department of Psychology, University of Notre Dame, Notre Dame, IN 46556; [k]Amsterdam Center for Language and Communication, University of Amsterdam, 1012 VT Amsterdam, The Netherlands; [l]Hertfordshire Community National Health System Trust, Welwyn Garden City AL7 1BW, United Kingdom; [m]Institute of Estonian and General Linguistics, University of Tartu, Tartu 50090, Estonia; [n]Faculty of Humanities/Logopedics, University of Oulu, 90014 Oulu, Finland; [o]Laboratoire sur le Langage, le Cerveau, et la Cognition, CNRS-Université de Lyon, 69675 Bron Cedex, France; [p]Department of English Philology, Ivane Javakhishvili Tbilisi State University, 0179 Tbilisi, Georgia; [q]Centre for General Linguistics, D-10117 Berlin, Germany; [r]Institute of Education, University of Leipzig, D-04109 Leipzig, Germany; [s]Department of Linguistics, University of Athens,15784 Ilissia, Greece; [t]Department of English Literature and Linguistics, Bar Ilan University, Ramat-Gan 52900, Israel; [u]Department of Psychology, University of Milano-Bicocca, 20126 Milan, Italy; [v]College of Law, Nihon University, Tokyo 101-8375, Japan; [w]Training Center for Foreign Languages and Diction, Tokyo University of the Arts, 110-8714 Tokyo, Japan; [x]Centre for Applied English Studies, University of Hong Kong, Hong Kong; [y]Department of Second Language Acquisition, University of Maryland, College Park, MD 20742; [z]Faculty 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; [ii]Department 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 [nn]Department 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

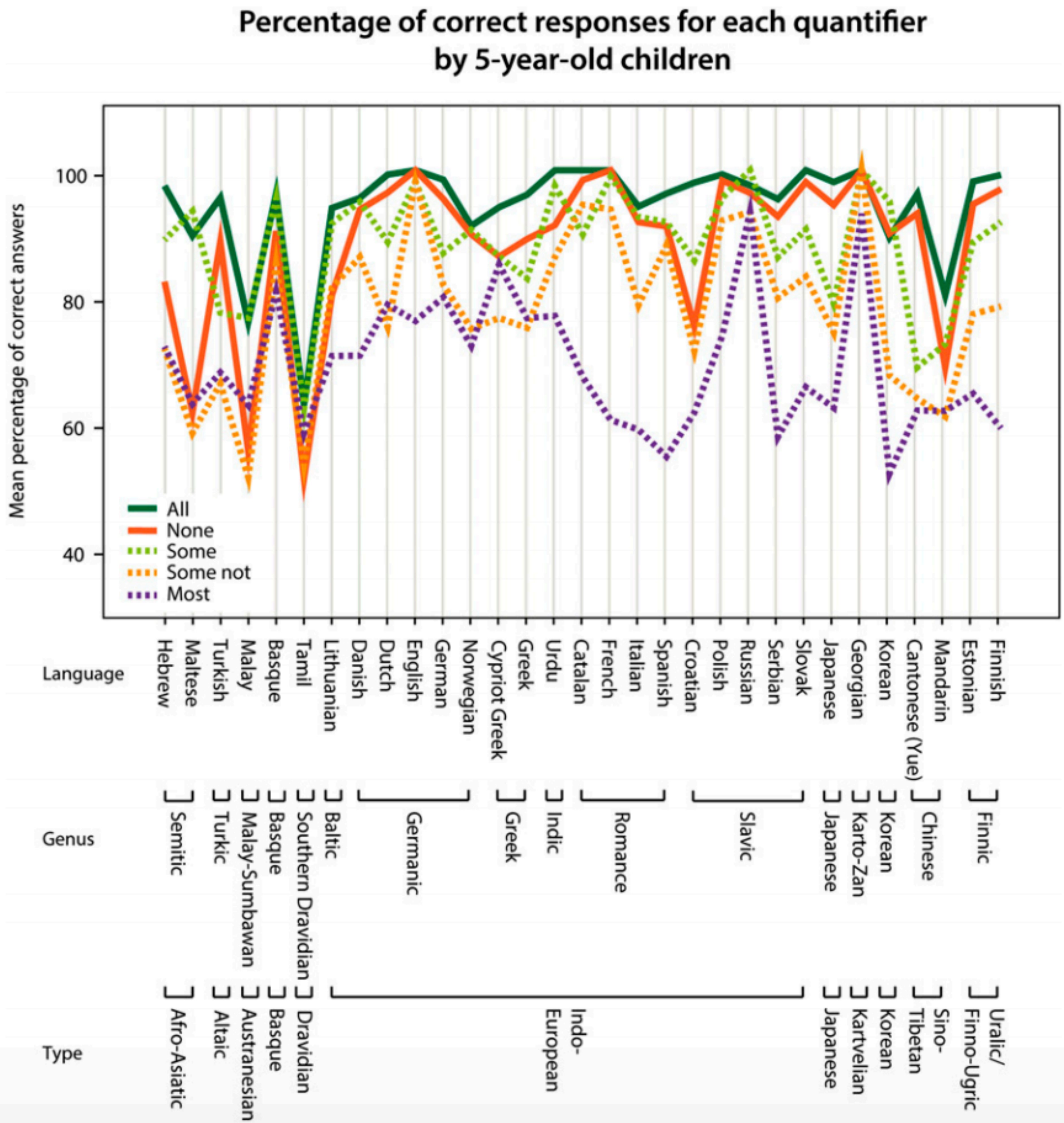
- 768 children (mean age = 5.5; range = 5.00–5.11; 398 female)
- 536 adults (293 female).
- 31 languages: Basque, Cantonese (Yue) Chinese, Catalan, Croatian, Cypriot Greek, Danish, Dutch, English, Estonian, Finnish, French, Georgian, German, Greek, Hebrew, Italian, Japanese, Korean, Lithuanian, Malay (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)



Katsos et al., 2016

Discussion

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

What do we know about complexity and 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

A R T I C L E I N F O

Article history:
Available online 7 March 2017

Keywords:
Generalized quantifiers
Semantic complexity
Corpus analysis
Generalized linear regression models
Analysis of deviance

A B S T R A C T

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.

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;*

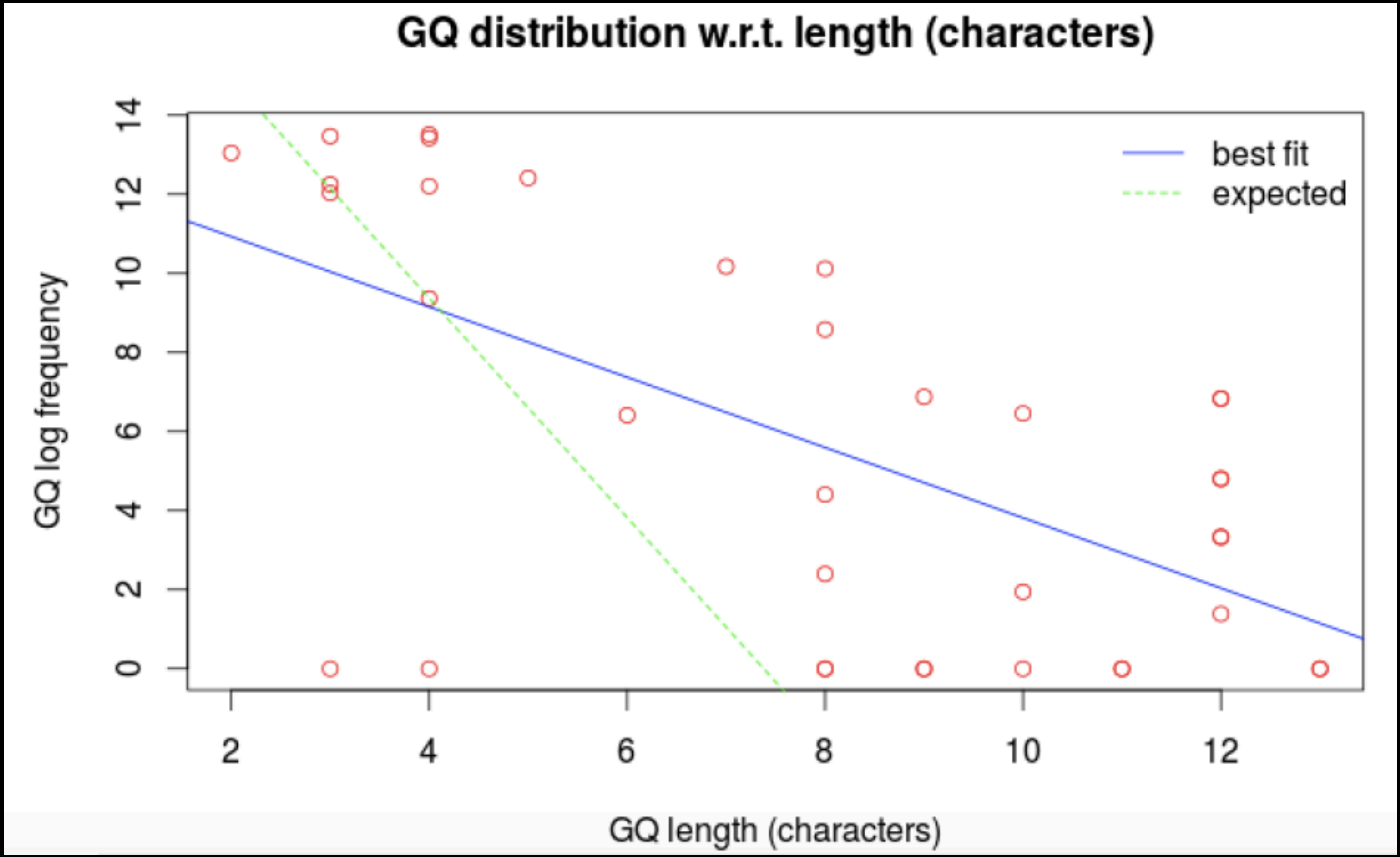
- *‘more/jjr than/in [a-z]{1,12}/cd’, viz., the same as before, but with ‘more’ a comparative adjective.*

- 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



Szymanik & Thorne, 2017

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

Lang Resources & Evaluation
https://doi.org/10.1007/s10579-022-09578-4

PROJECT NOTES

The semantically annotated corpus of Polish quantificational expressions

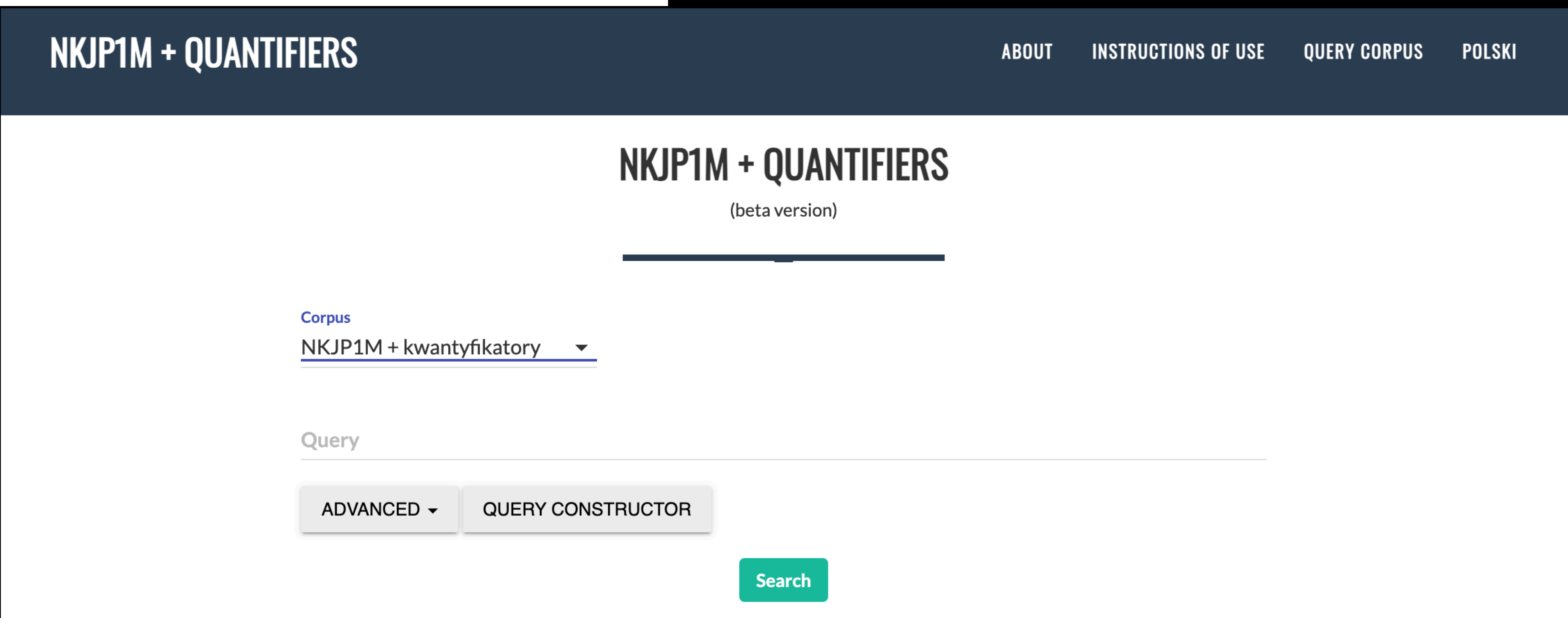
Jakub Szymanik[1] · Witold Kieraś[2]

Accepted: 11 January 2022

Abstract The paper presents a manually annotated corpus of Polish quantificational expressions. The quantifier annotation was conducted on top of existing gold-standard data for Polish as its separate layer. This paper releases the data and gives an overview of the corpus and related tools. As far as we know, this is the first large-scale annotation of generalized quantifiers together with their crucial semantic properties, including monotonicity profile. We also discuss the potential further use of the corpus in linguistics and cognitive science.

Keywords Generalized quantifiers · Corpus · Annotation

kilka = between 2 and 10
kilkanaście = between 11 and19
kilkadziesiąt = between 20 and 99
kilkaset = between 100 and 999



RaRa and the Universals Archive

Universal 1199:

Posted in Universals Archive

Universal 1199:

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.

Keywords
: quantification, generalized quantifier, NP

Domain
: syntax, semantics

Type
: unconditional

Status
: achronic

Quality
: absolute

Basis
: unspecified

Source
: **Barwise & Cooper 1981**: 177, U1

Counterexamples
: 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 all, many, two 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).

Universal 1203:

Posted in Universals Archive

Universal 1203:

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.

Keywords

  quantification, monotonicity, NP

Domain

  syntax, semantics

Type

  unconditional

Status

  achronic

Quality

  absolute

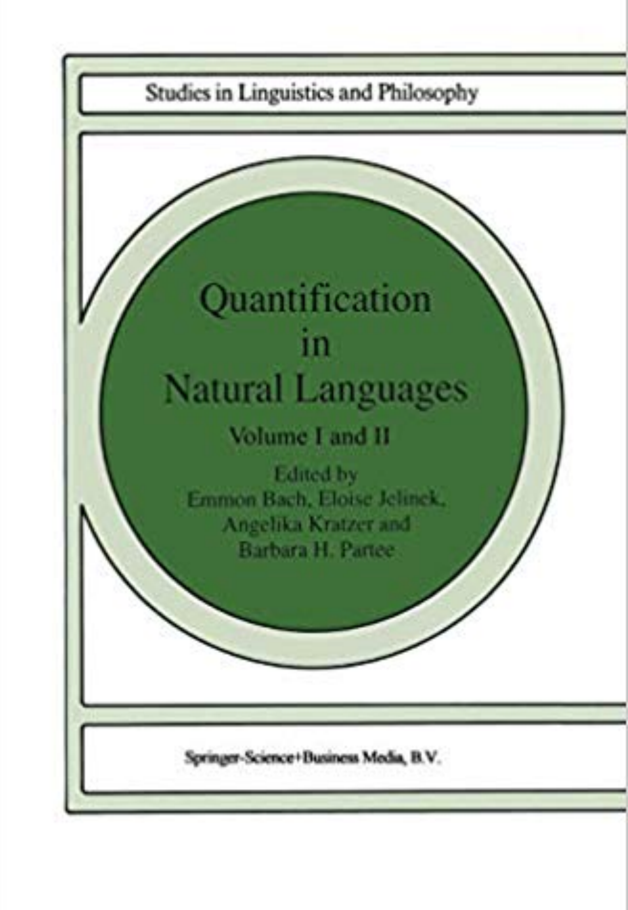
Basis

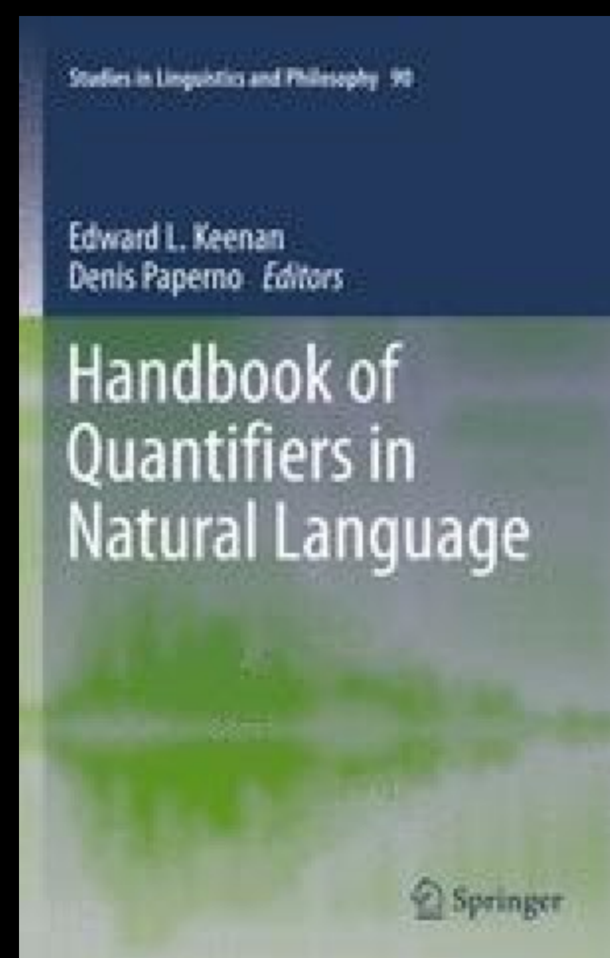
  unspecified

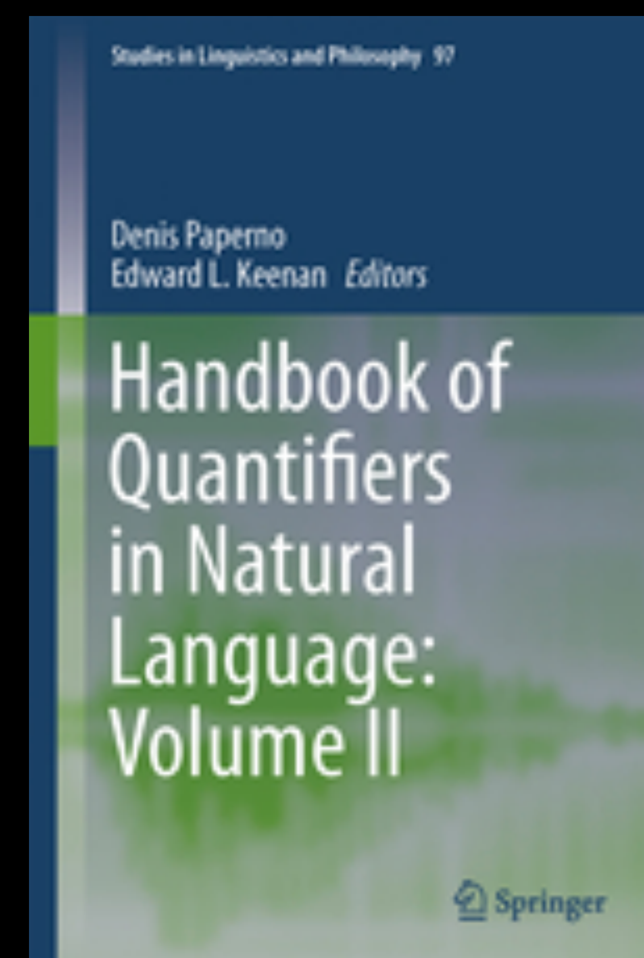
Source

  Barwise & Cooper 1981: 187, U6

Counterexamples


Studies in Linguistics and Philosophy
Quantification
in
Natural Languages
Volume I and II
Edited by
Emmon Bach, Eloise Jelinek,
Angelika Kratzer and
Barbara H. Partee
Springer-Science+Business Media, B.V.


Studies in Linguistics and Philosophy 90
Edward L. Keenan
Denis Paperno Editors
Handbook of
Quantifiers in
Natural Language
Springer


Studies in Linguistics and Philosophy 97
Denis Paperno
Edward L. Keenan Editors
Handbook of
Quantifiers
in Natural
Language:
Volume II
Springer

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 allows to formulate universal constraints

2: GQ theory enables natural computational representations and complexity measures

3: ‘Older’ tools in formal learning do not explain universals

4: Neural learnability explains many universals, across domains [plus connections to evolution and complexity]

5: The developmental and cross-linguistic pictures are not very complete

New methodologies can shed light on “central” linguistic questions!

Combining simulations and experiments is necessary to answer the question why languages are what they are.

Outlook

- Scaling up experiments
- Are there different types of linguistic universals? E.g.: Cons vs Mon, Ext, Isom
- Unification of / more fundamental explanations of universals? Complexity vs learnability vs efficiency?
- Cross-linguistic work

Thanks!

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

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