Investigating the emergence of overspecification in an Iterated Learning setup

Jonas Nölle (Aarhus University)
Peeter Tinits (Tallinn University)
Stefan Hartmann (University of Hamburg)
Michael Pleyer (Universität Heidelberg)

IACS, Lublin, 20.06.2016

AARHUS UNIVERSITET

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

TALLINN UNIVERSITY

Outline

- Redundancy in sign systems
- Overspecification
- Coordination game experiments
  - Language pragmatics
  - Language dynamics
- Our study
- Conclusions

Redundancy

- Most natural semiotic systems have elements in them that can be considered redundant:
  i.e. on perfect conditions, the system would work equally well without them.

Redundancy

- Most natural semiotic systems have redundancy

```
(1) THE ROOM WAS NOT VERY LIGHT A SMALL OBLONG
(2) ----ROO------NOT-V-----I------SM----OBL----
(1) READING LAMP ON THE DESK SHED GLOW ON
(2) REA----------O------D----SHED-GLO--O--
(1) POLISHED WOOD BUT LESS ON THE SHABBY RED CARPET
(2) P-L-S-----O---BU--L-S--O------SH-----RE--C------
```

(Shannon 1951: 54)

- Some estimates
  - Written English – ~50% (Shannon 1951; Newman & Gerstman 1950; Garner & Carson 1960)
  - Written Hungarian – poetry 40%, newspapers 67%, two young girls 71% (Fomagy 1961)

Why is there redundancy?

- Redundancy can be useful:
  - Noisy environments
  - Inattentive listeners
- But commonly speakers want to strive for efficiency and reduce redundancy whenever possible, languages are in balance between different pressures
- How does redundancy emerge, and how is it maintained in these signalling systems?

(e.g. Zipf 1935; Martinet 1950; Beckner et al. 2008; Kirby et al. 2015).

Overspecification

- „Languages differ essentially in what they *must* convey and not in what they *may* convey. “ Jakobson, 1959

- Overspecification – redundancy of meaningful elements
- Occurs if semantic distinctions are marked when redundant or irrelevant for the current interaction.

Overspecification

Karok:

- *pa:ɸ -kírih* ‘ throw into fire’
- *pa:ɸ -kúrih* ‘throw into water’
- *pa:ɸ -rúprih* ‘throw in through a solid’

(McWhorter 2007)

English:

- Where is the IACS conference this year?
- **It** take**s** place in Lublin.

Overspecification

Questions

- Why do we have overspecification in some places, but not in others?
- How does overspecification come about?

Experimental approach

- Small semiotic experiments
  - Simple tasks to coordinate a meaning (help pick the right object) or describe some object.
  - With natural language or with a taught artificial language
  - Alone, with an imagined addressee, or someone who is with you

FOOTBALL

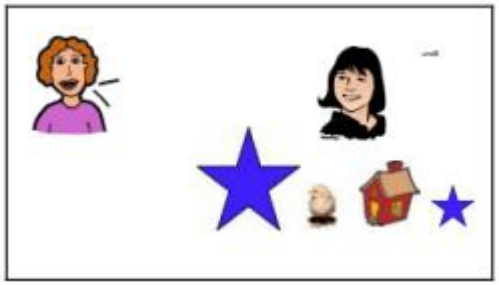
TUPIN

ICE SKATER

(Clark & Wilkes-Gibbs 1986; Kirby, Cornish, Smith 2008)

Earlier research

- Experimental pragmatics
  - Give the participant(s) a task to solve with natural language.
  - Possible to test for various conditions or use, or short term changes in customs.
- For example tests for overspecification:
  - „Pass me the small (blue) star" or „pass me the (blue) star"

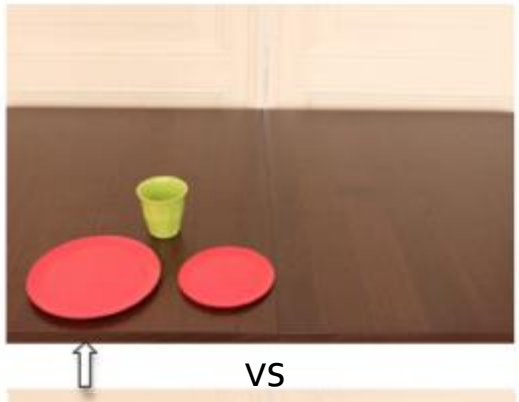


(Davies & Katsos 2009)

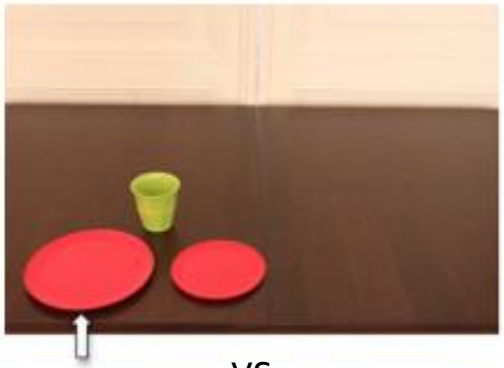
Overspecification experiments

- Using Dutch language to describe the target object

- Less OS:



vs



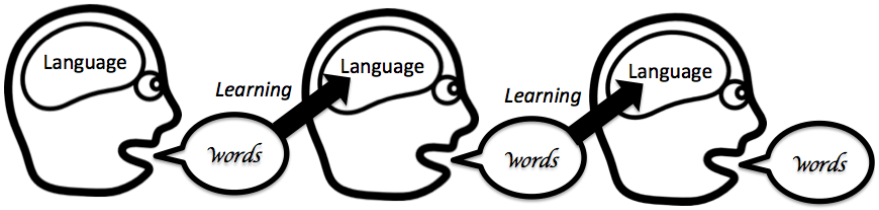
- More OS:

vs

(Koolen et al. 2015)

Earlier research

- Iterated learning
  - Teach participant a language, give them a task that requires linguistic output, and give that output to the next participant.

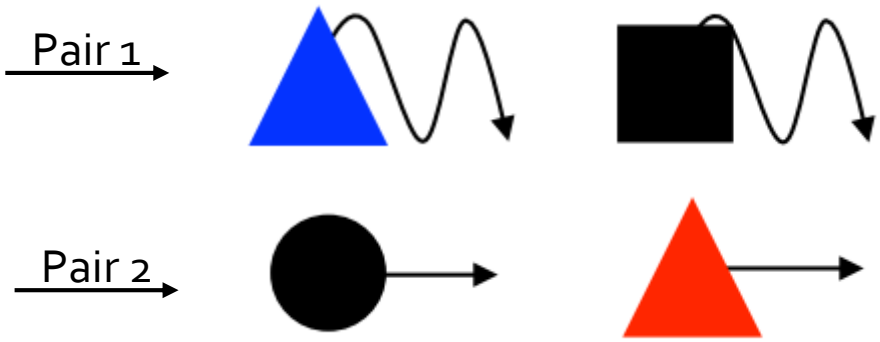


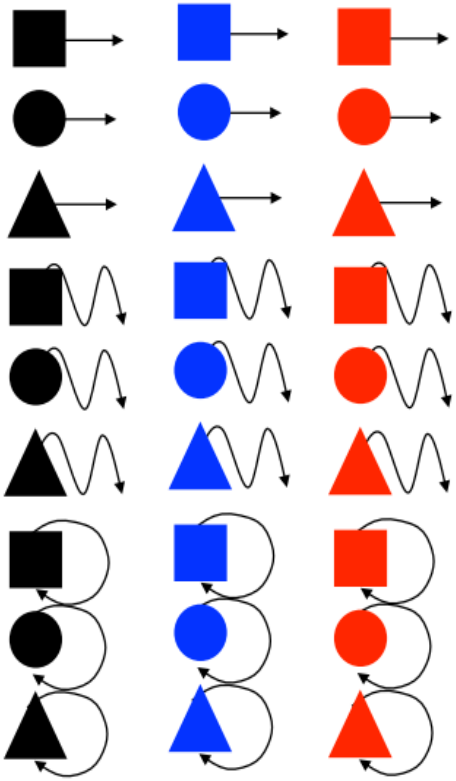
  - Within experimental time cultural transmission, and repeated learning and use can transform the language.

(Kirby et al. 2008; Kirby et al. 2014)

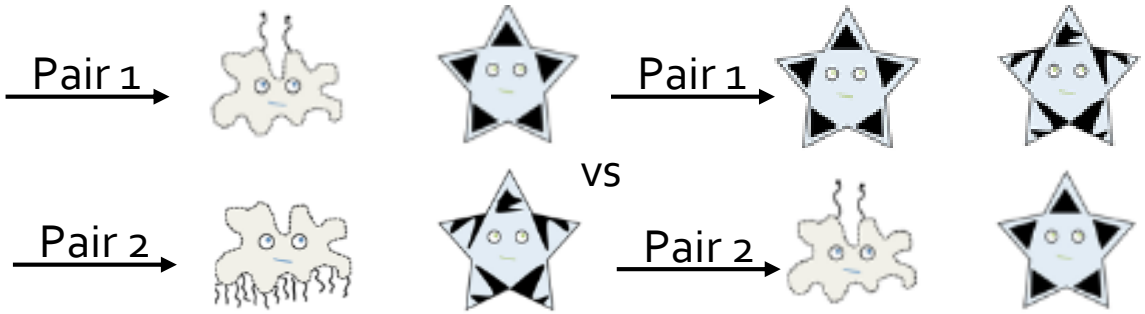
Iterated learning

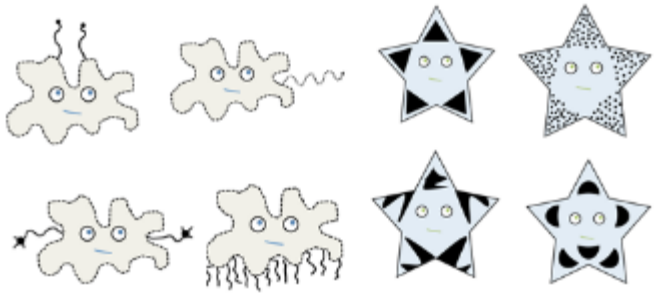
- Silvey et al. 2014





- Winters et al. 2015





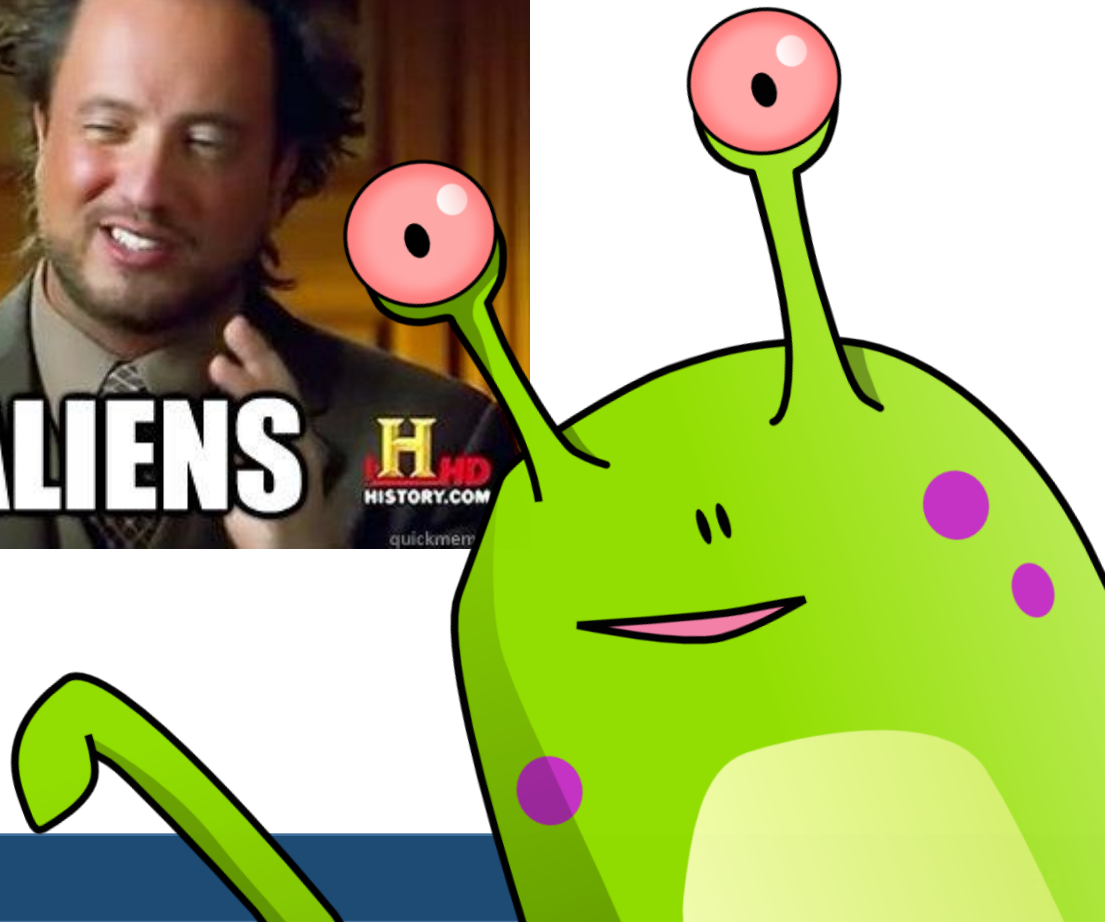
(Kirby et al. 2008; Silvey et al. 2014; Winters et al. 2015)

The question

- Can this kind of contextual variation work as a mechanism for the emergence of overspecification?

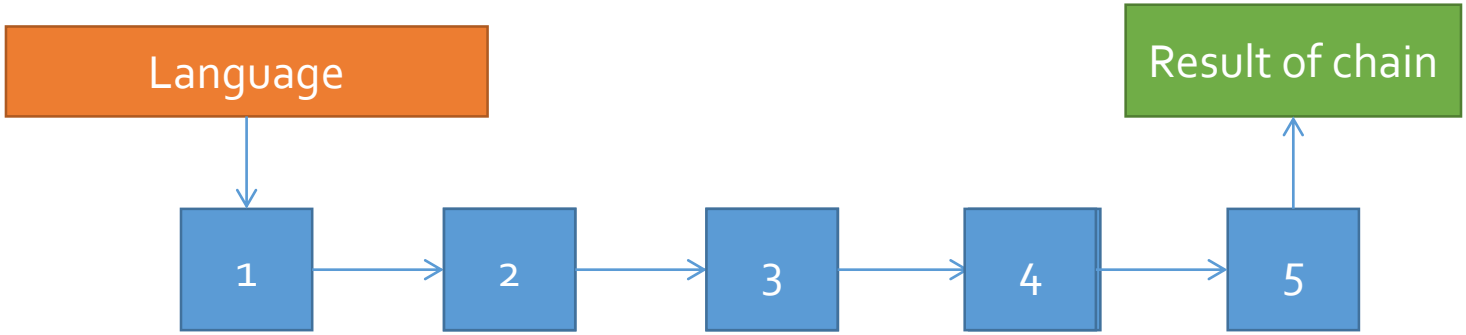
- An experiment
  - Artificial language
  - Iterated learning
  - Variation in overspecification
- > Differences between contexts


ALIENS
HISTORY.COM



Methods and participants

- Web-based study "Learn an alien language!"
- 205 volunteers recruited via social media (92f, 100m, 13*)
- 18 x 2 chains, 5 participants each (5 chains excl.)
- ~15 minutes per participant



Language

- Simple language with a non-obligatory color marker for 1st generation

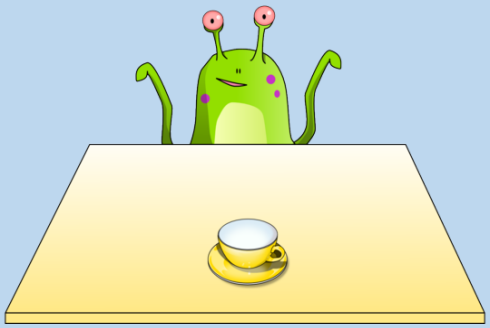
No contrast	Yellow	Blue
meeb	meeb li	meeb pu
fop	fop li	fop pu
kur	kur li	kur pu
yan	yan li	yan pu

Allows for very simple variation to use „li“ & „pu“ on 0 to 16 no-contrast items.

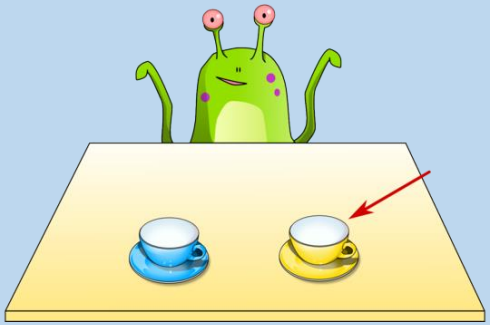
8 x li	8 x pu		

Experimental setup

Isolated Condition

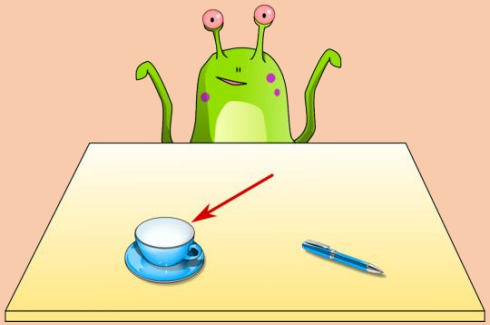


fop

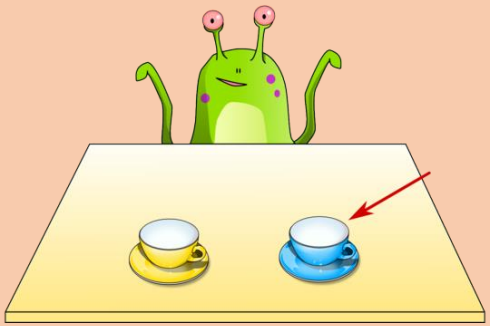


fop li

Composite Condition



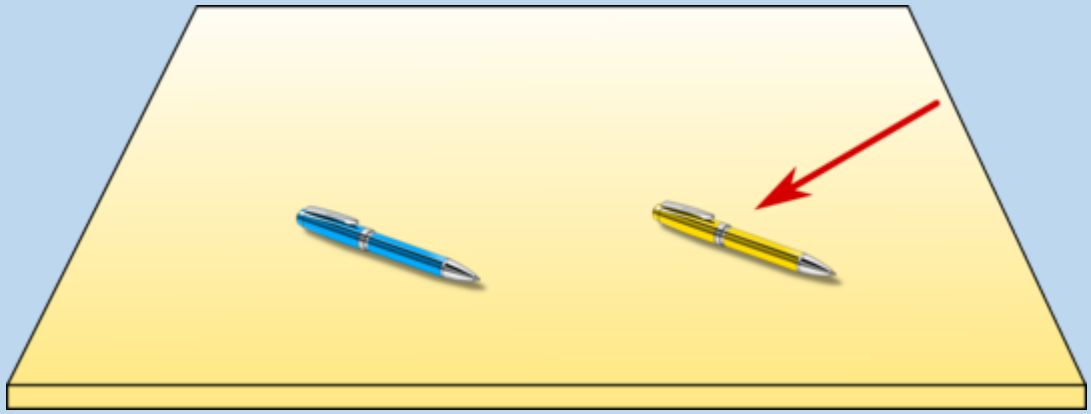
fop



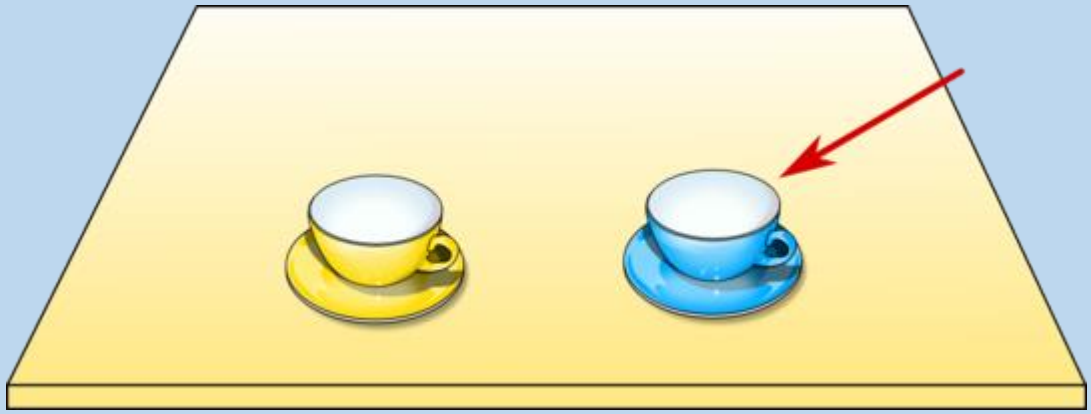
fop pu

Isolated context
(color irrelevant)

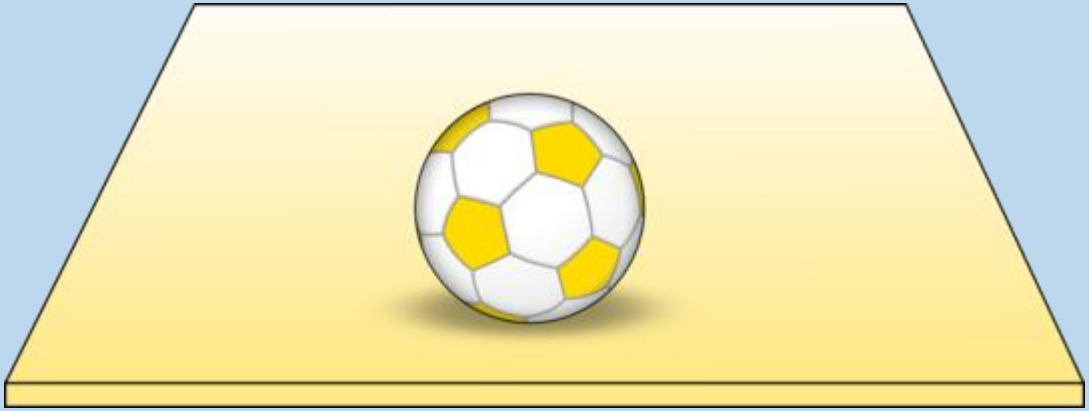
Distinction context
(color relevant)



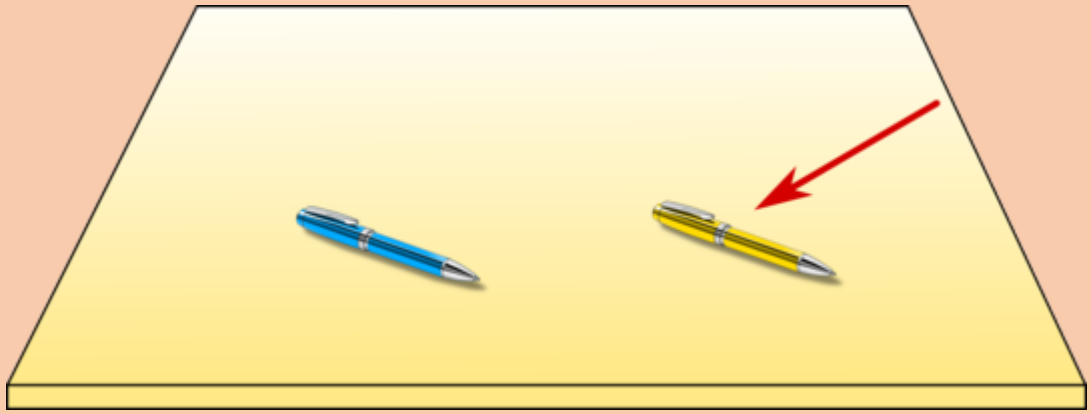
Distinction context
(color relevant)



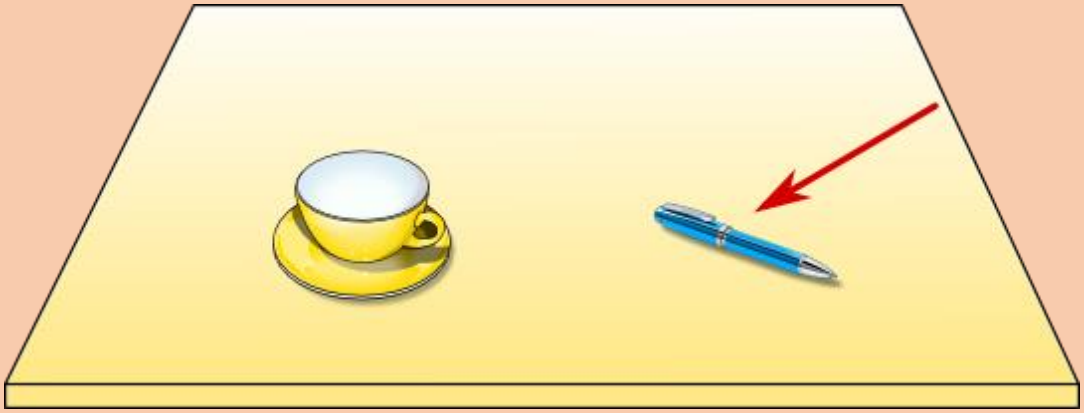
Isolated context
(color irrelevant)



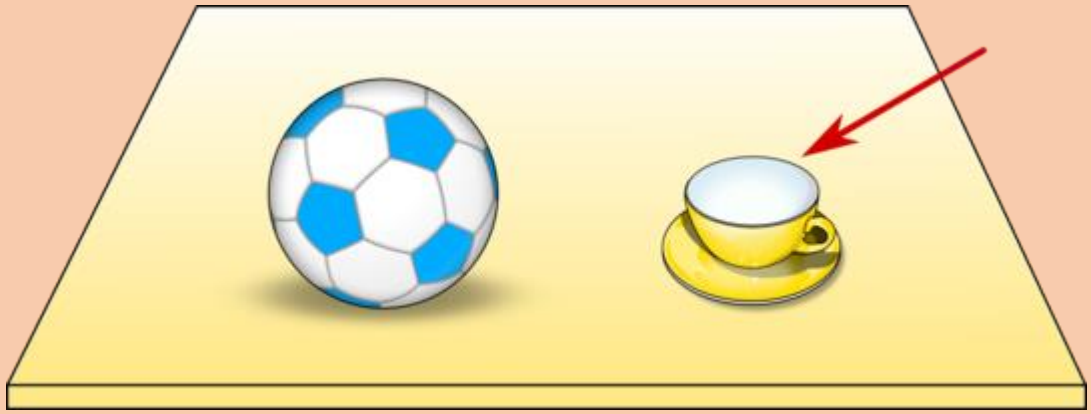
Distinction context
(color relevant)



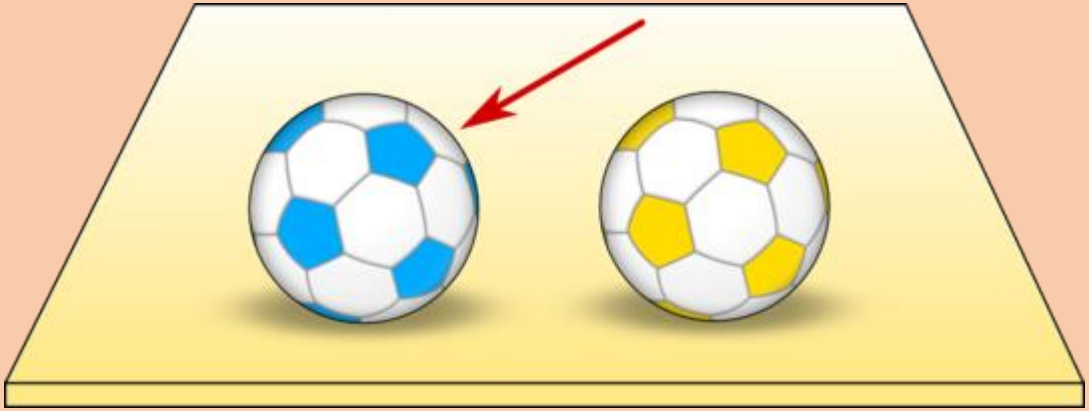
Composite context
(color irrelevant)



Composite context
(color irrelevant)

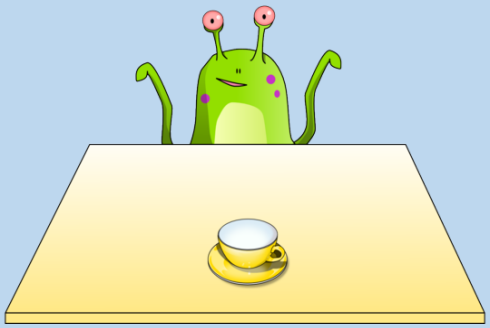


Distinction context
(color relevant)

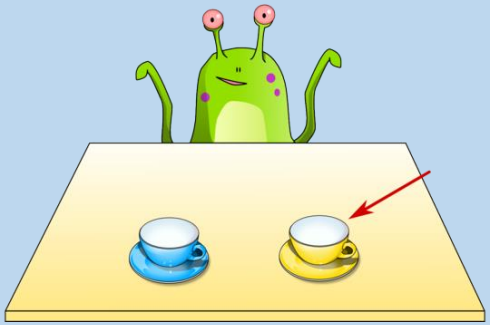


Experimental setup

Isolated Condition

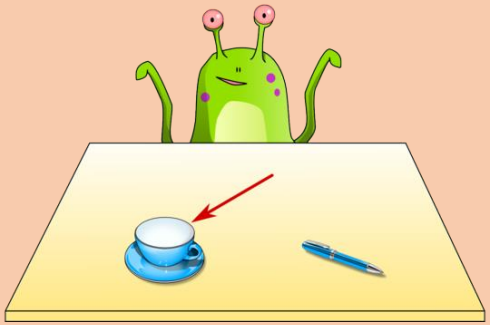


fop

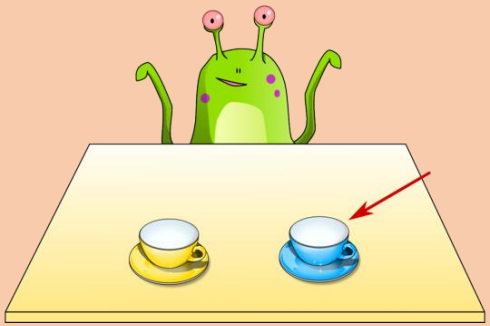


fop li

Composite Condition

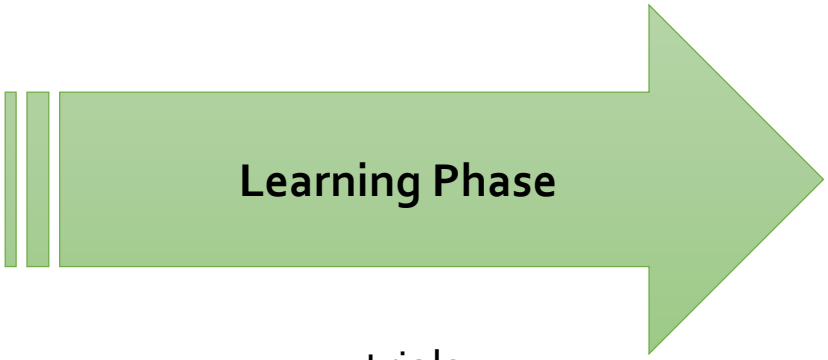


fop



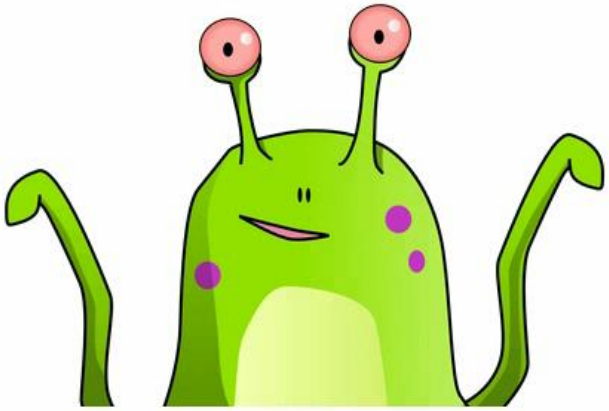
fop pu

Experimental setup



32 trials

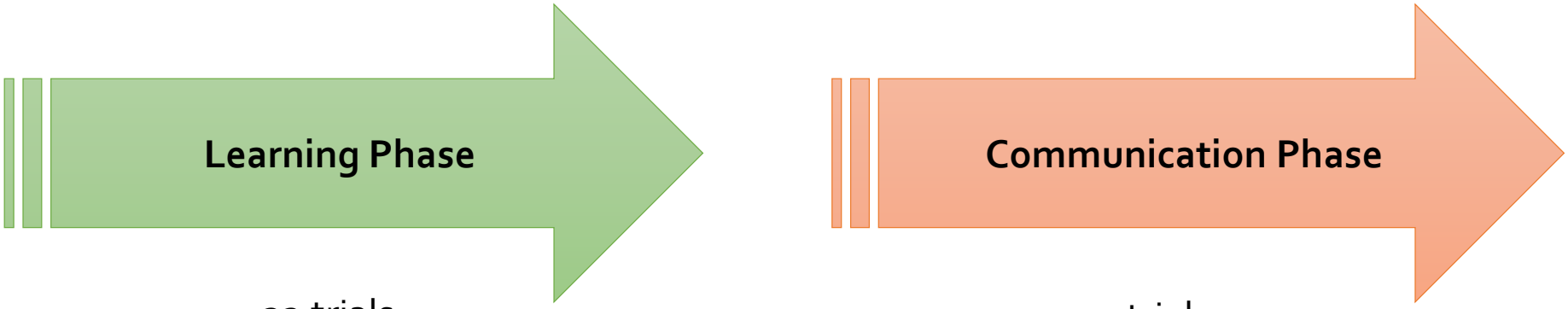
Welcome to our little experiment! It will take about 15 minutes.



This is Ari, an alien from the planet Tsamtrah. Over the next few minutes you will learn to communicate with Ari. You are dealing with a shipment of earth items to Ari which got mixed up on space travel. Ari has restored most of the order, but needs your help in the finishing steps.

Your task is to help Ari in various situations where Ari is unsure which object to pick

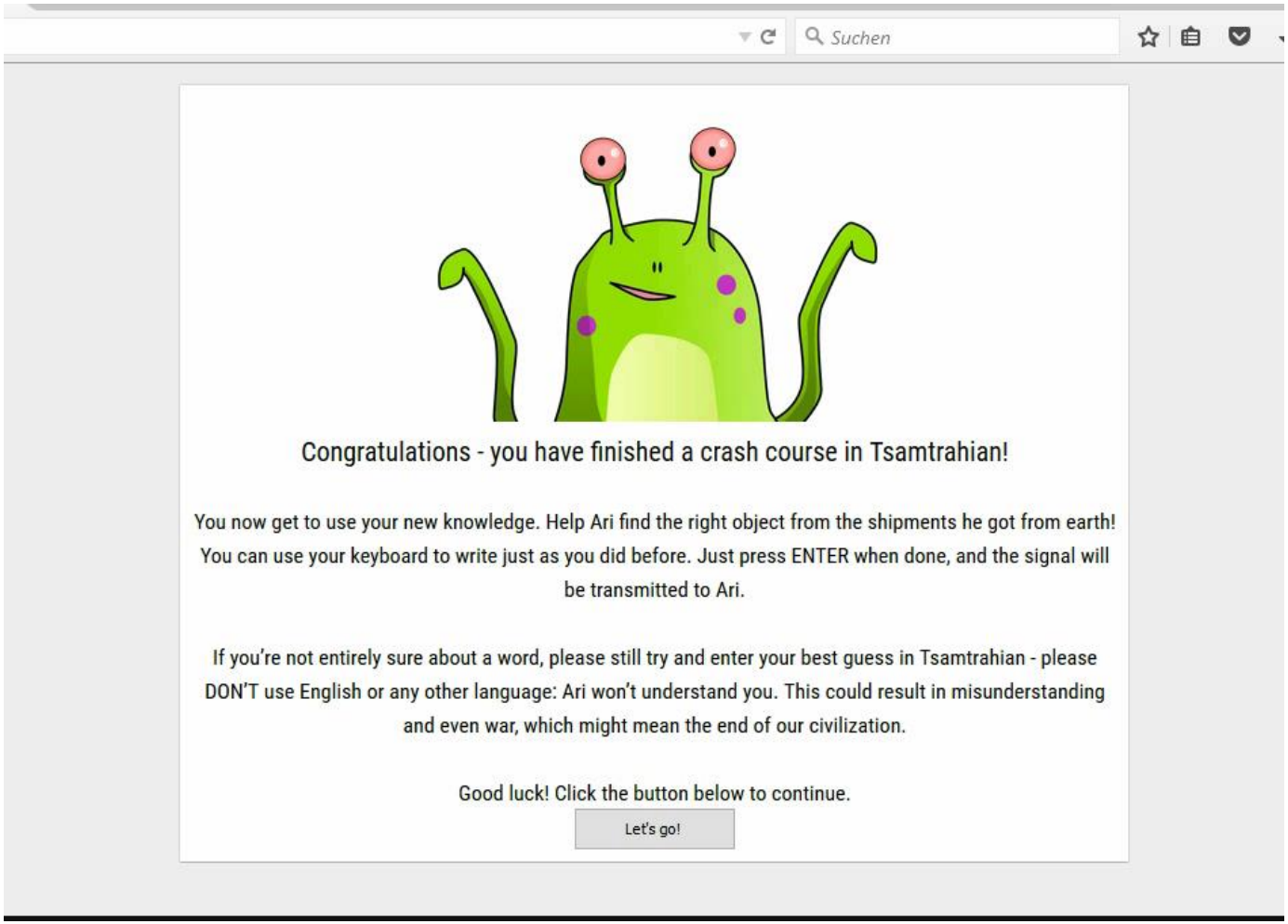
Experimental setup



32 trials

32 trials

+
Questionnaire


Suchen
Congratulations - you have finished a crash course in Tsamtrahian!
You now get to use your new knowledge. Help Ari find the right object from the shipments he got from earth! You can use your keyboard to write just as you did before. Just press ENTER when done, and the signal will be transmitted to Ari.
If you're not entirely sure about a word, please still try and enter your best guess in Tsamtrahian - please DON'T use English or any other language: Ari won't understand you. This could result in misunderstanding and even war, which might mean the end of our civilization.
Good luck! Click the button below to continue.
Let's go!

Results

Chain from the Composite condition

Results

Chain from the Isolated condition

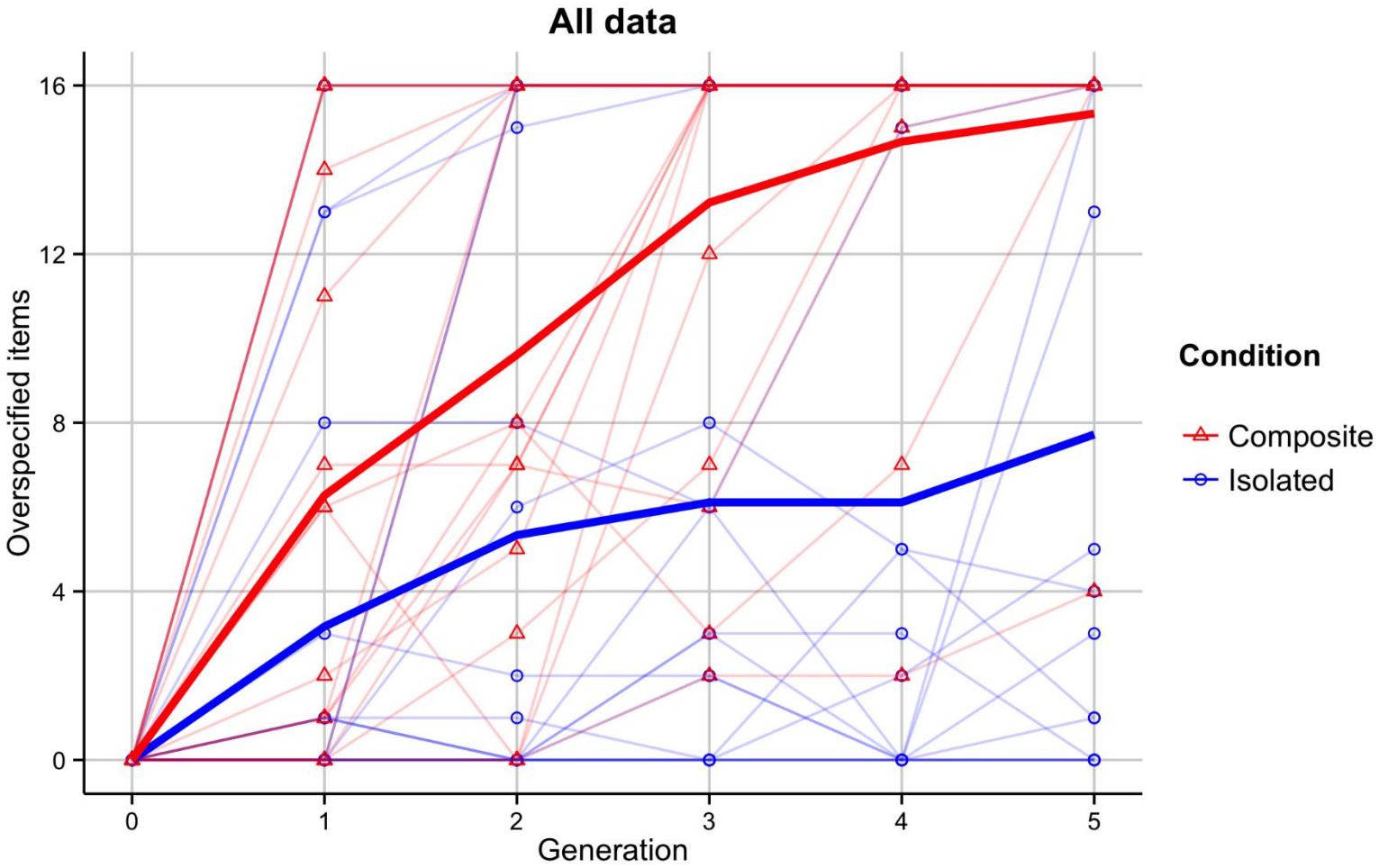
Results

meeb
fop
kur li
yan
yan

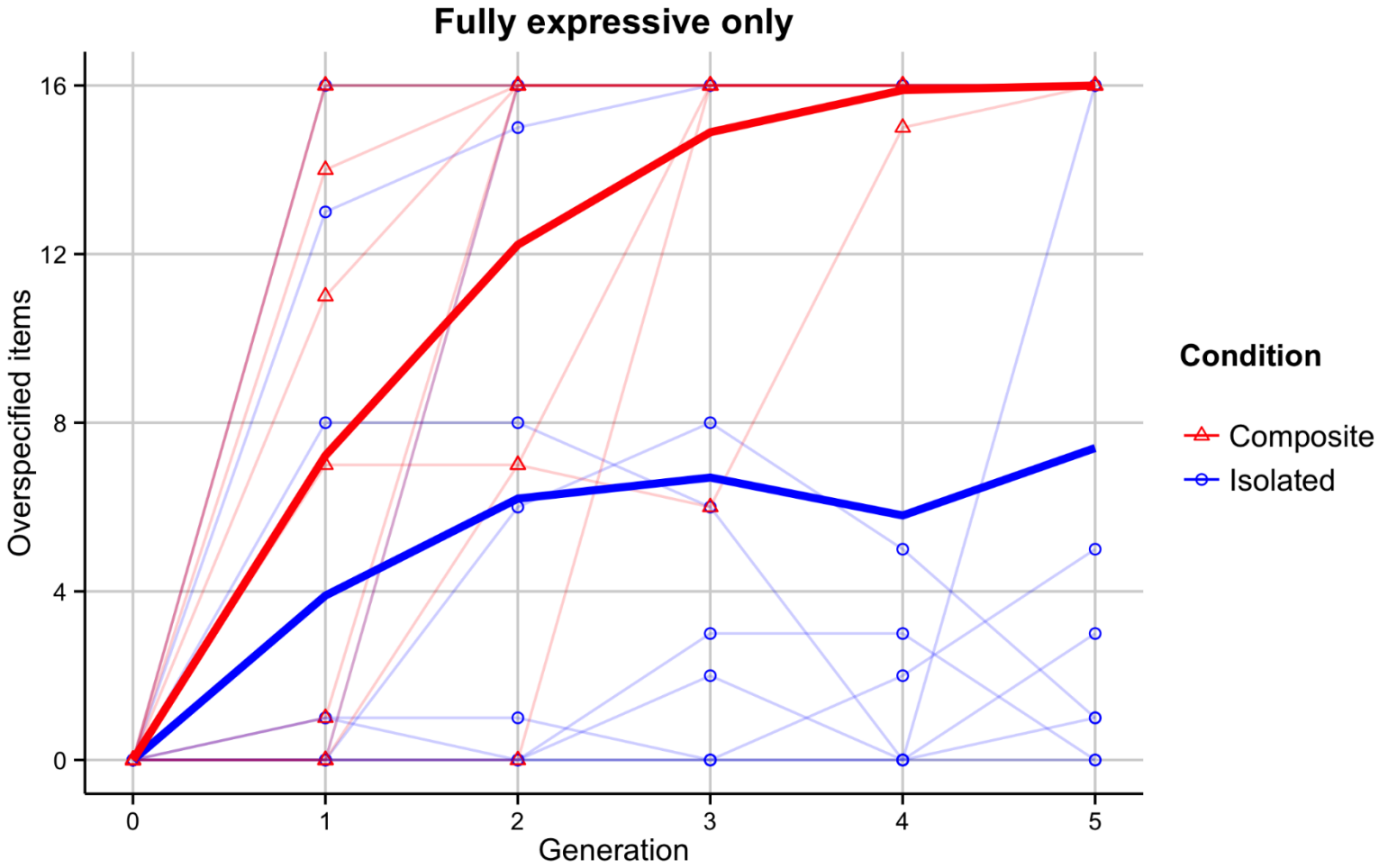
Results

- Analysis of fully expressive chains:
  - 20 of the 41 chains still expressed 4 distinct root words and 2 distinct color terms in the final generation
  - Homonymy: In three chains contrast was lost between two root words
  - Unmarking: In 6 chains one (or more) root word was lost leading always to obligatory marking of color
  - In 7 chains the color distinction was lost over time

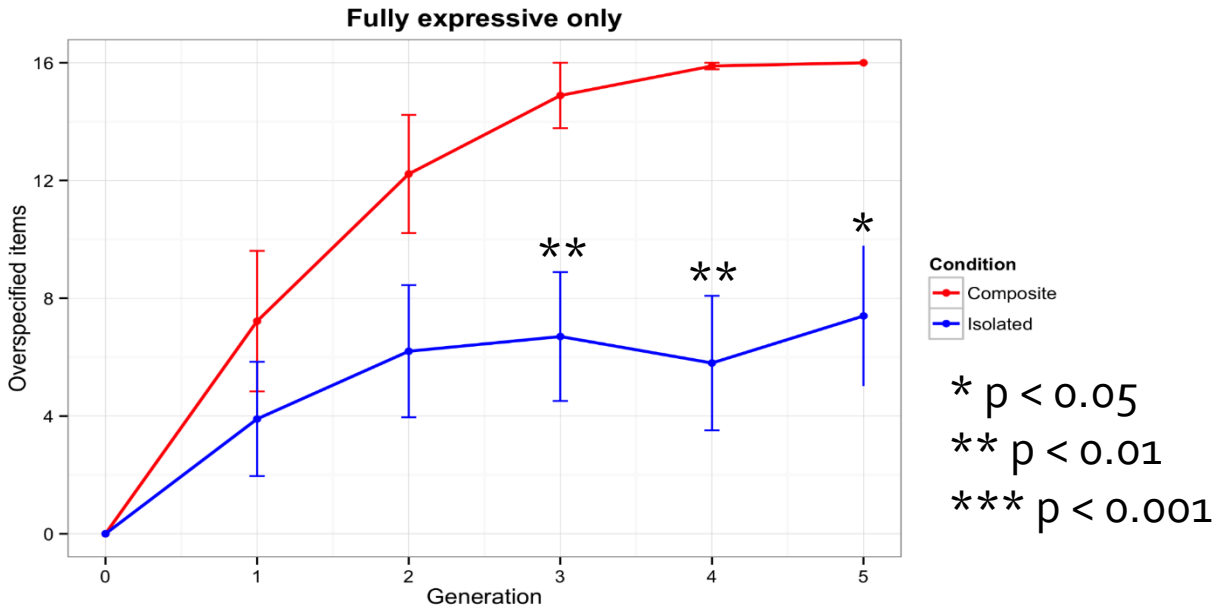
Results


All data
Overspecified items
16
12
8
4
0
0
1
2
3
4
5
Generation
Condition
Composite
Isolated

Results


Fully expressive only
Overspecified items
16
12
8
4
0
0
1
2
3
4
5
Generation
Condition
Composite
Isolated

Resulting languages



* p < 0.05
** p < 0.01
*** p < 0.001

- Increase in overspecification in both conditions, but significantly more (and faster) in the composite condition
- Difference in generation 5:
  - all chains: Mann-Whitney U-Test, U=69, p<0.001, Cliff's δ =0.6
  - fully systematic chains only: U=18, p<0.01, Cliff's δ =0.6

Discussion

What does it tell us?

- Contextual pressures shape the evolution of overspecification in artificial language learning and use

- In some contexts, overspecification appears more efficient to the user

- Implications for typological diversity?

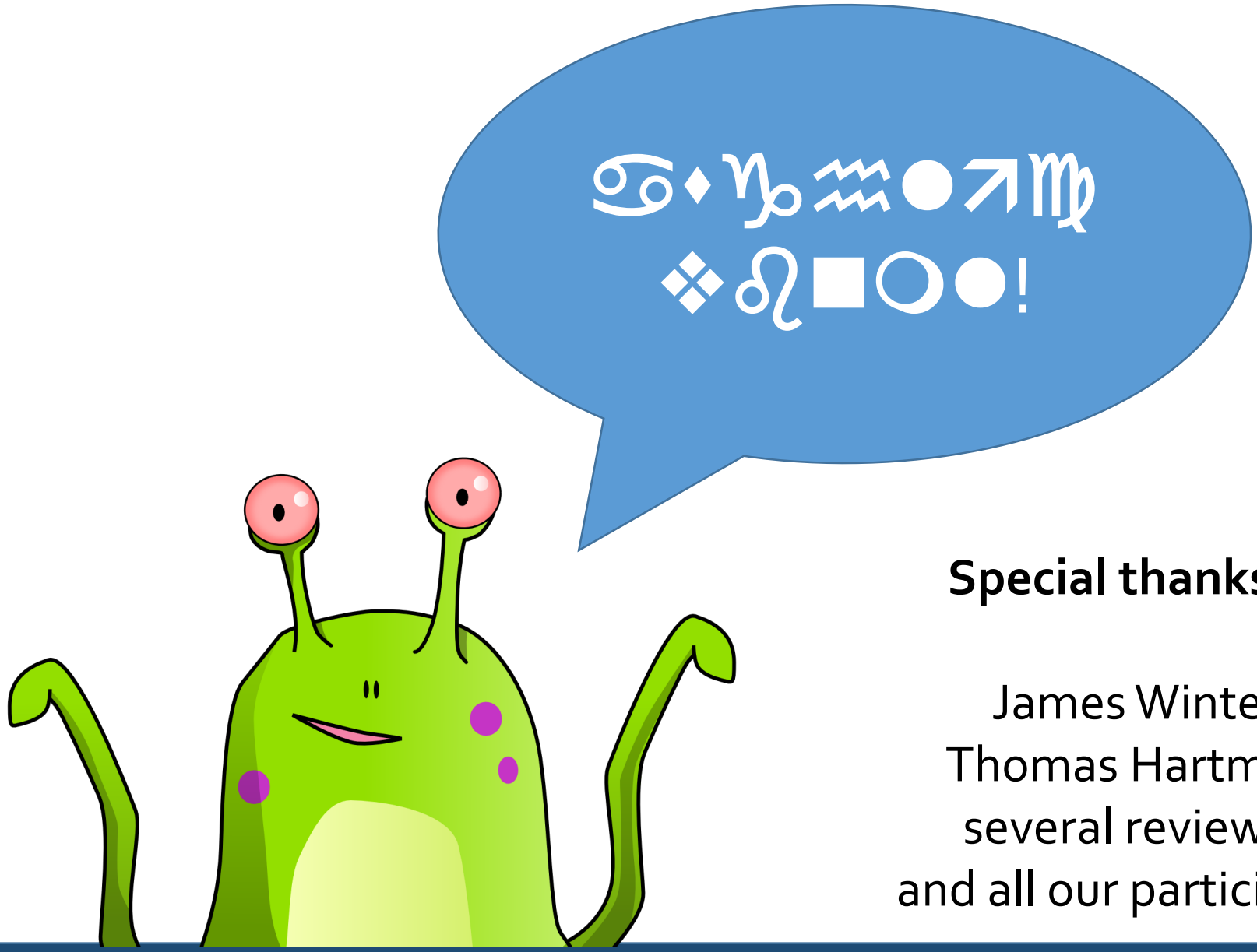
What's next? Open questions and ideas for further research

Linguistic Niche Hypothesis

- Can the differences in situational context employed in our study be linked up with different situations of language use?

Communication games

- Follow-up studies using a communication game setup with interactions between real participants could extend the domain of relevance.



Special thanks to:

James Winters
Thomas Hartmann
several reviewers
and all our participants!

References 1

- Brighton, H., Smith, K., & Kirby, S. (2005). *Language as an evolutionary system*. *Physics of Life Reviews*, 2, 177–226.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., Schoenemann, T.) (2009). Language is a complex adaptive system. Position paper, *Language Learning, 59*, Supplement 1, 1–27.
- Christiansen, M. H. & Chater, N. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5):489-509.
- Dahl, Östen 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.
- Fusaroli, R. & Tylén, K. (2012) Carving language as social coordination, Interaction studies, 13(1), 103-124.
- Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. *Frontiers in human neuroscience*, *5*(February), 11. doi:10.3389/fnhum.2011.00011
- Kam, C. L. H., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(1650), 151–195.
- Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: when learners change languages. *Cognitive psychology*, *59*(1), 30–66. doi:10.1016/j.cogpsych.2009.01.001
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(31), 10681–6. doi:10.1073/pnas.0707835105
- Kusters, Wouter. 2003. *Linguistic Complexity: the Influence of Social Change on Verbal Inflection*. PhD Dissertation, University of Leiden.
- Ladd, D. R., Roberts, S. G., & Dediu, D. (2015). Correlational studies in typological and historical linguistics.Annual Review of Linguistics, 1, 221-241. doi:10.1146/annurev-linguist-030514-124819.
- Lupyan G, & Dale R (2010). Language structure is partly determined by social structure.*PloS one, 5* (1) PMID: 20098492

References 2

- Reali, F., Chater, N. & Christiansen, M.H. (2014). The paradox of linguistic complexity and community size. In E.A. Cartmill, S. Roberts, H. Lyn & H. Cornish (Eds.): The Evolution of Language: Proceedings of 10th international conference (pp 270-277). Singapore: World Scientific
- Roberts, S & James Winters (2012). Social Structure and Language Structure: the New Nomothetic Approach. *Psycology of Language Learning, 16* (2), 89-112 :10.2478/v10057-012-0008-6
- Roberts S, Winters J (2013) Linguistic Diversity and Traffic Accidents: Lessons from Statistical Studies of Cultural Traits. PLoS ONE 8(8): e70902. doi:10.1371/journal.pone.0070902
- Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116(3), 444–9. doi:10.1016/j.cognition.2010.06.004
- Trudgill, P. (2012). On the sociolinguistic typology of linguistic complexity loss. In F. Seifart, G. Haig, N. P. Himmelmann, D. Jung, A. Margetts, and P. Trilsbeek (Eds.) *Potentials of Language Documentation: Methods, Analyses, and Utilization* (pp. 96-104). Honolulu: University of Hawai'i Press
- McWhorter, J. (2007). *Language Interrupted*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195309805.001.0001
- Nichols, Johanna 2009. Linguistic complexity: A comprehensive definition and survey. In Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language complexity as an evolving variable*, 110-125. Oxford: Oxford University Press.
- Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language complexity as an evolving variable*, 98-109. Oxford: Oxford University Press.
- Thurston, W.R. (1987). *Processes of change in the languages of north-western New Britain.* Canberra: Pacific Linguistics B-99
- Tily, H., & Jaeger, T. F. (2011). Complementing quantitative typology with behavioral approaches: Evidence for typological universals. *Linguistic Typology*, 1–12.
- Tinits, P (2014)Language stability and morphological complexity in situations of language contact. An experimental paradigm. *Conference proceedings for the International Congress of Linguists 19*
- Wray, A. ; Grace, G. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*, 543–578.