

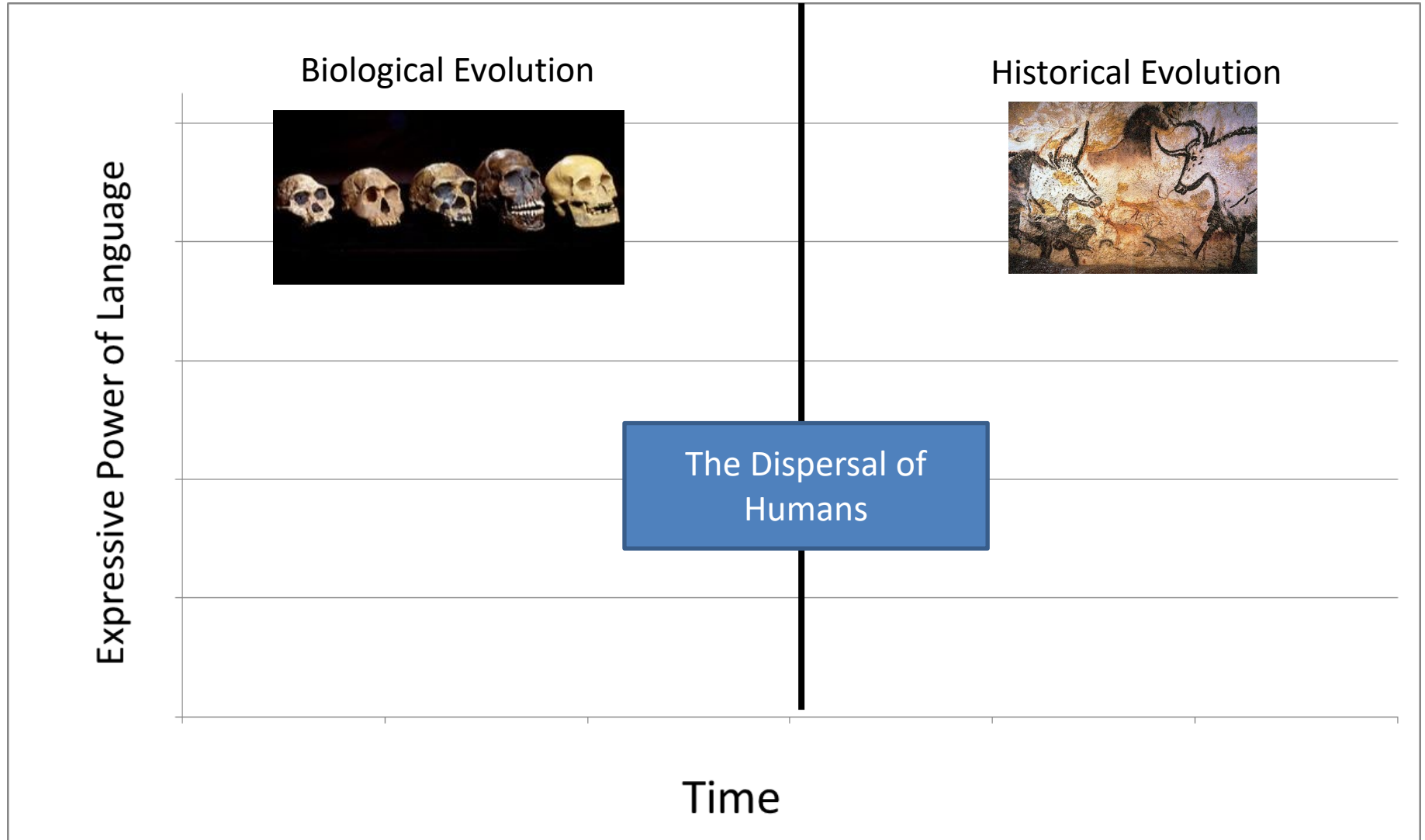
What language creation in the deaf tells us about the human capacity for language.

Jesse Snedeker
City University of Hong Kong
December 8, 2025

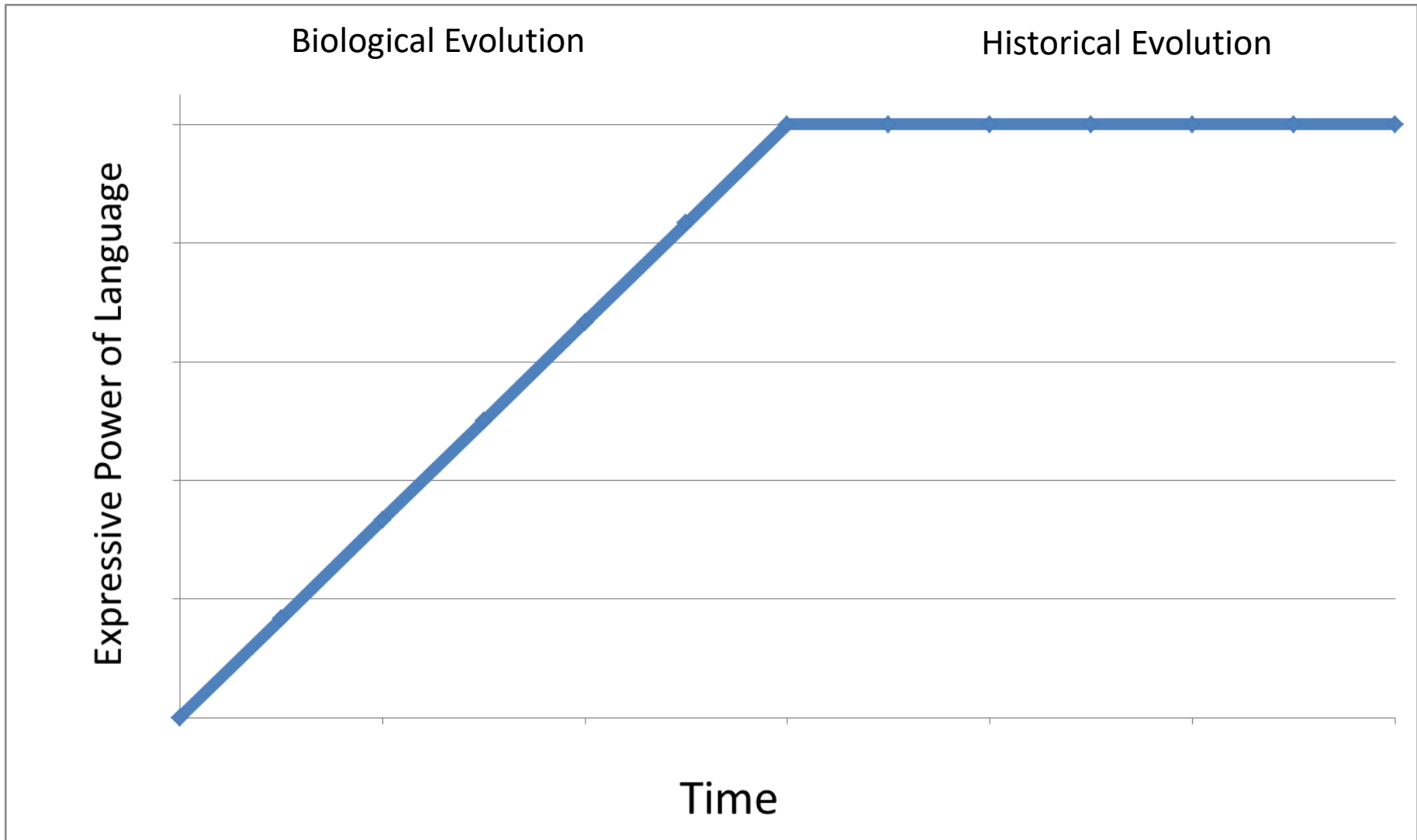




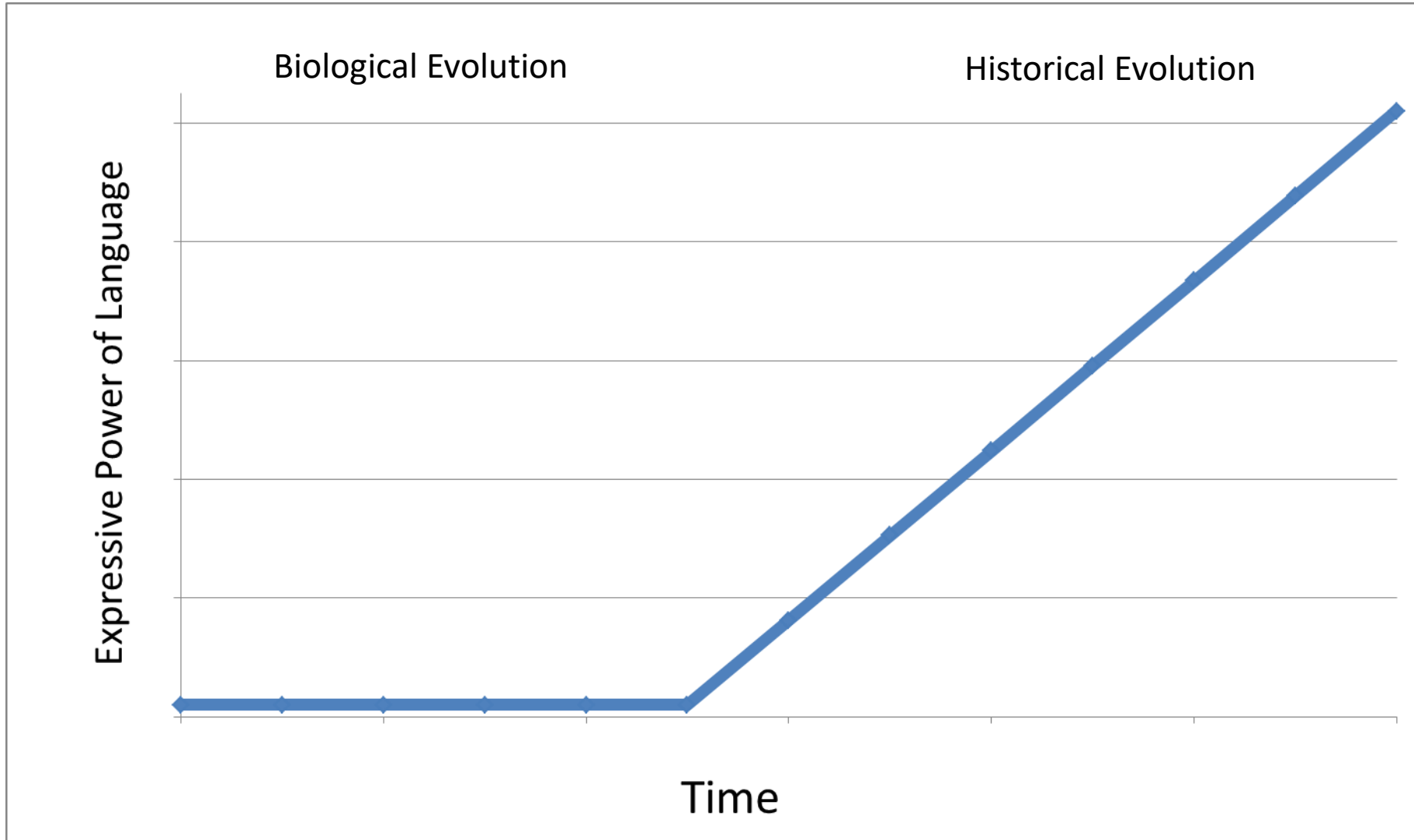
Where does language come from?



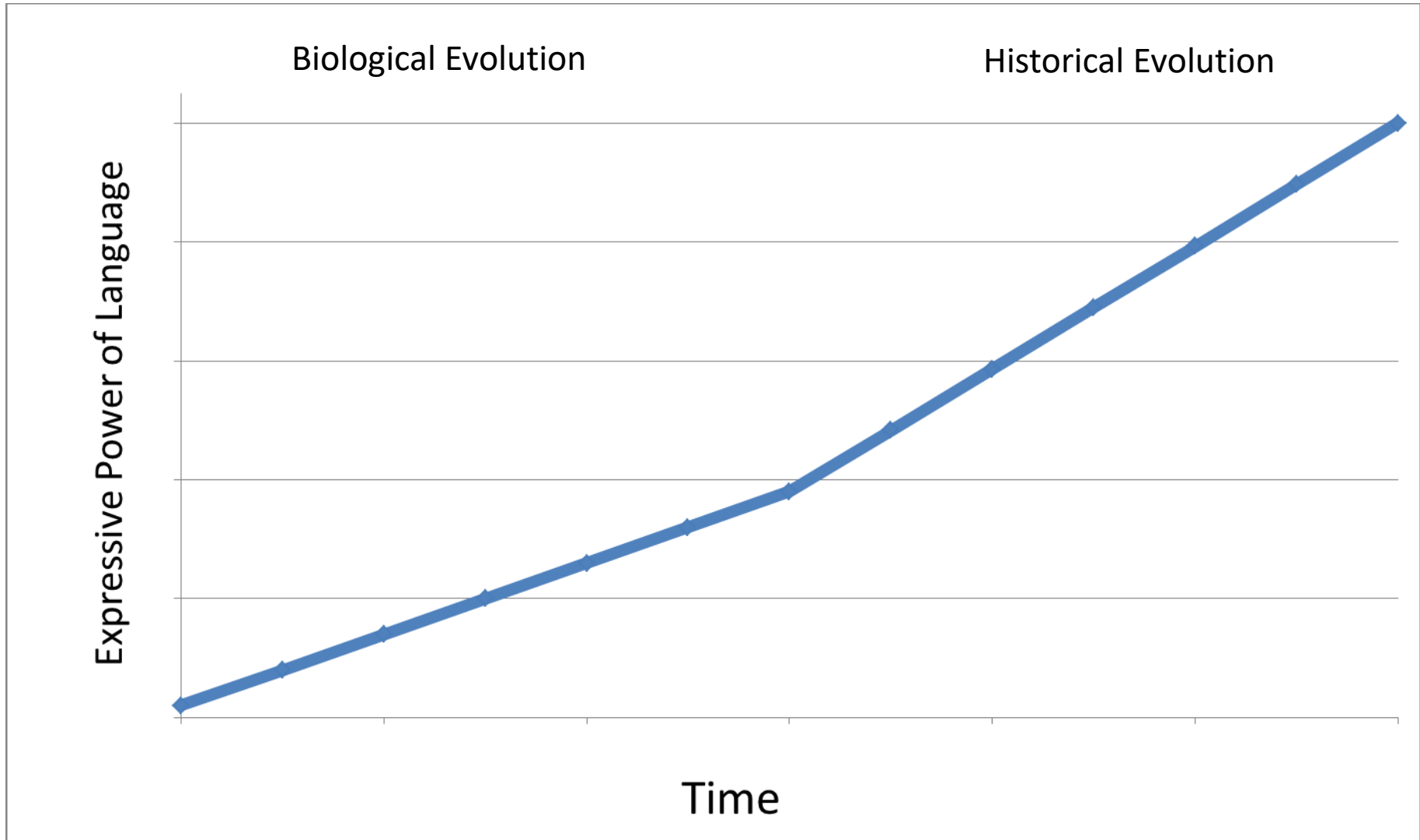
Biological Evolution



Historical Evolution



Both History and Biology



Problem: no record of the first 50,000 years



Solution: language creation in deaf communities



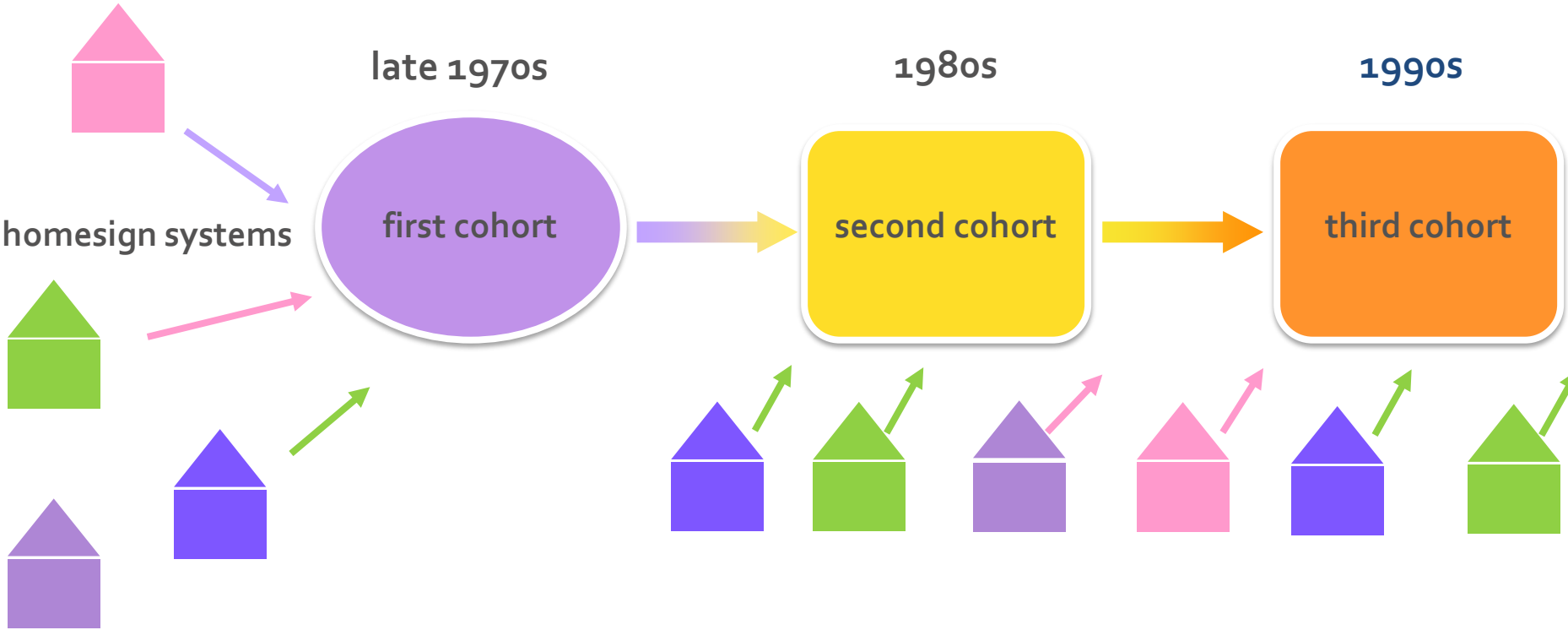
The primary force behind this work:



Annemarie Kocab
Johns Hopkins



History of Nicaraguan Sign Language



Children create these languages with little outside influence

- Minimal contact with existing sign languages
- Parents of homesigners produce less complex gestures than child
- Fail to understand child
- Emerging languages do not share features with community language
 - Ex: Spanish is SVO, NSL is verb final



Roadmap

Case Studies

- Quantifiers
- Recursion
- Agents and Patients

Theoretical conclusions

- Language is shaped by evolutionary constraints
- That may be largely conceptual in nature

A. Quantifiers



Annemarie Kocab and Kathryn Davidson

Quantifiers

- Express relations between two sets
 - Some (dogs) (bite people)
 - All (humans) (are mortal)
 - No (child) (enjoys being told what to do)
- Create chaos at the syntax-semantics interface
- Semantically complex
- But argued to be universal

Where do quantificational concepts come from?

- From learning a natural language?
 - If so, they might take many generations to emerge
- From the structure of the human mind?
 - If so, they may emerge rapidly in a new language
- Elicitation study with 28 NSL users from the first 3 cohorts

Existential:

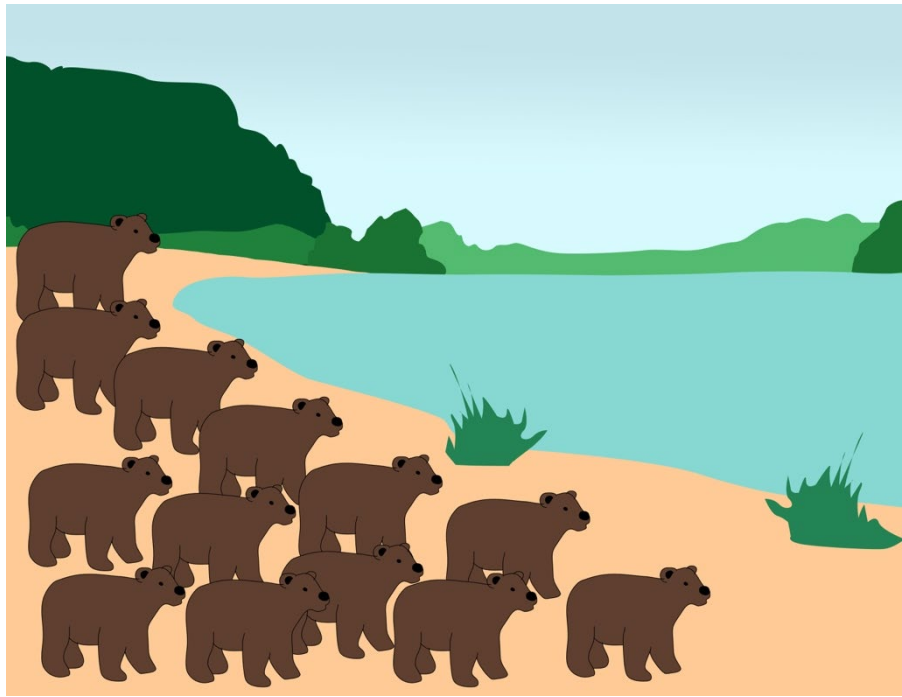
For some X , property P is true.
(many, some, a few)

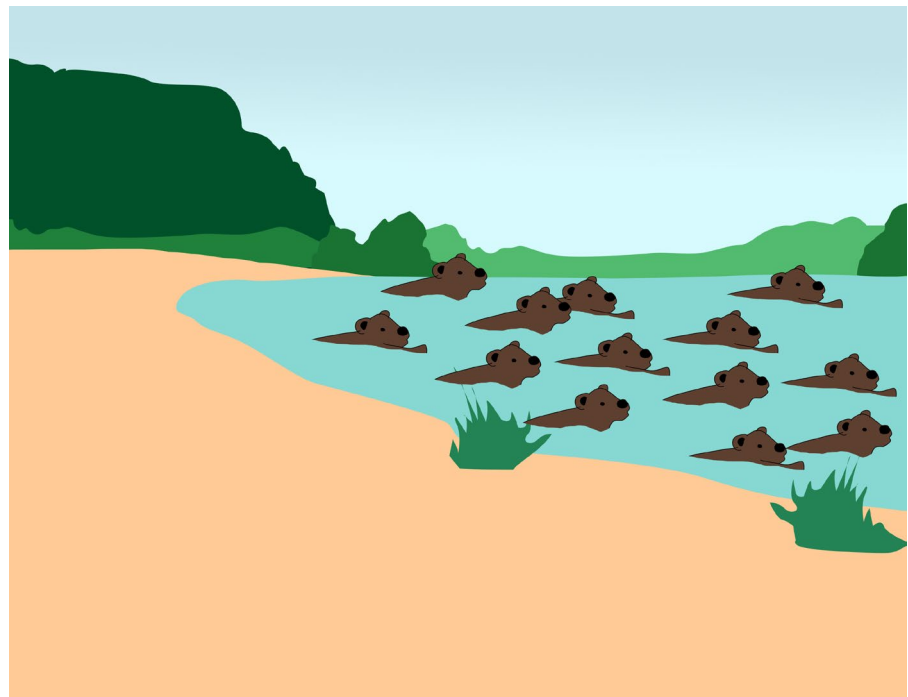
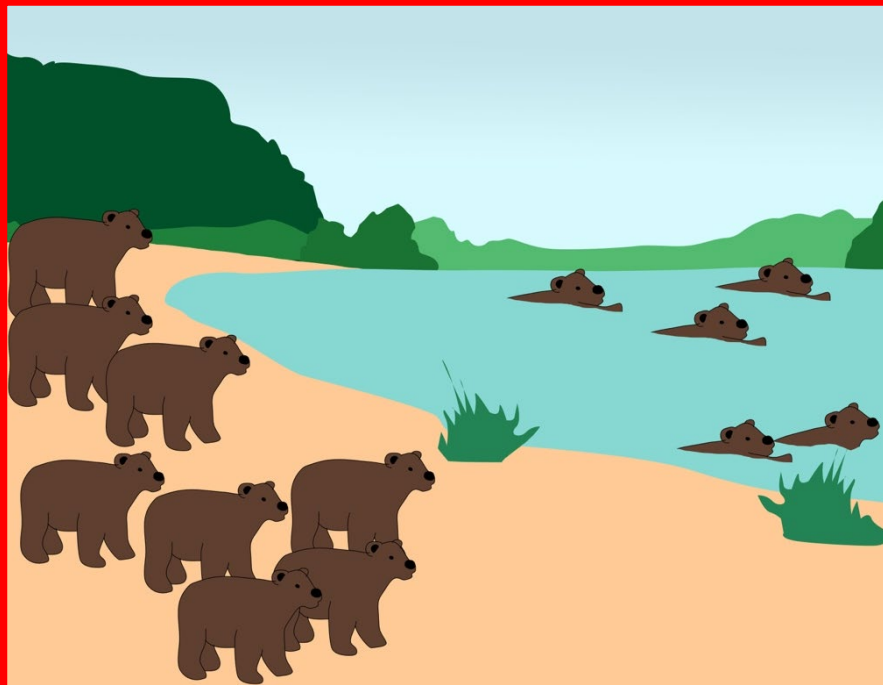
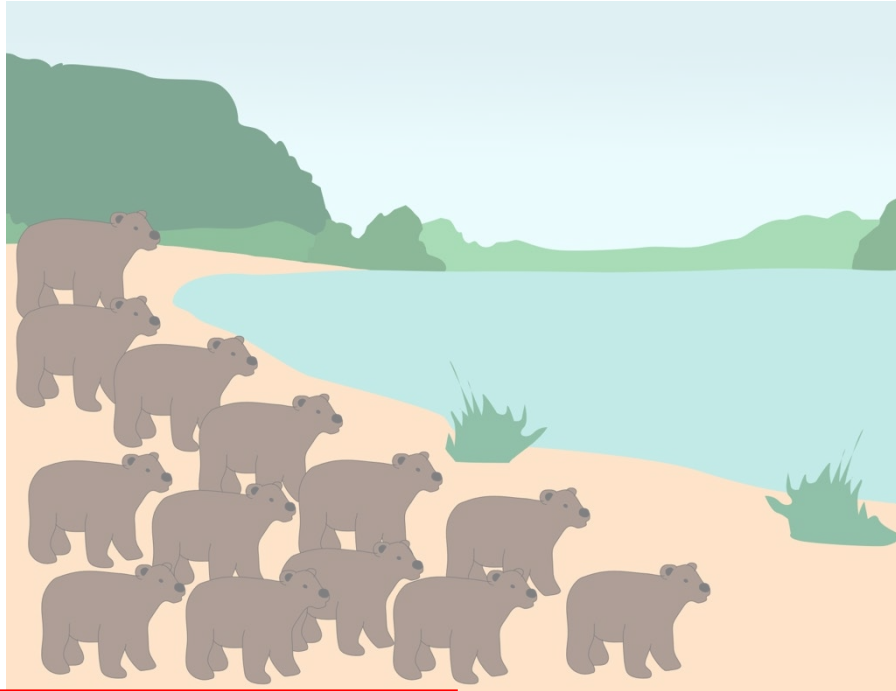
Negative:

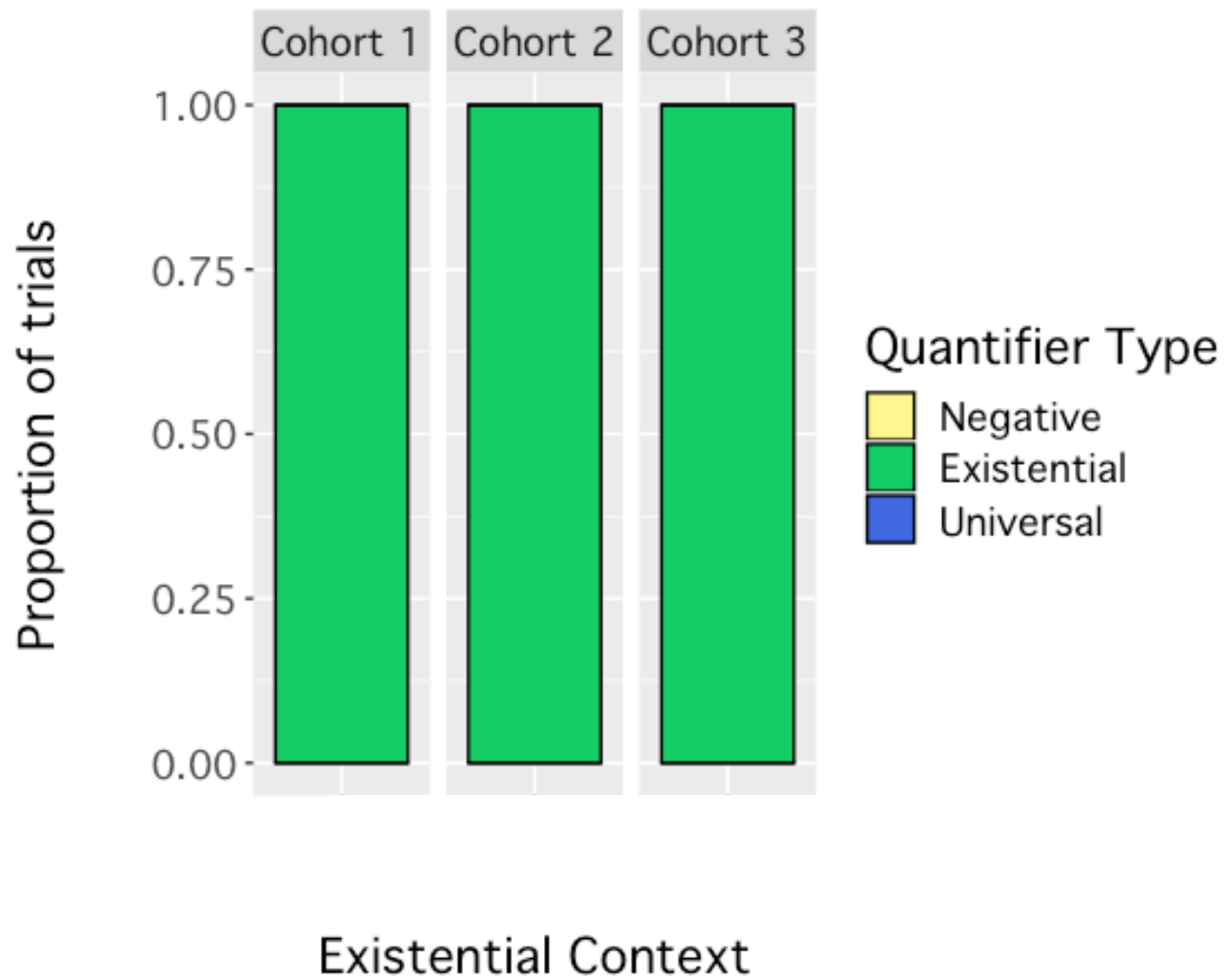
For no X , property P is true.
(no, none)

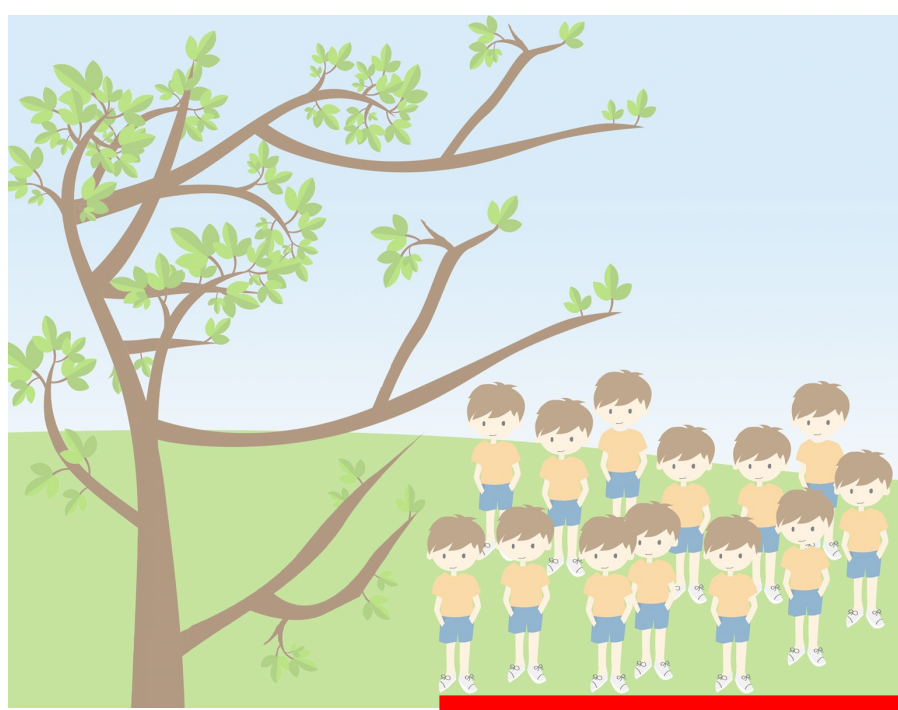
Universal:

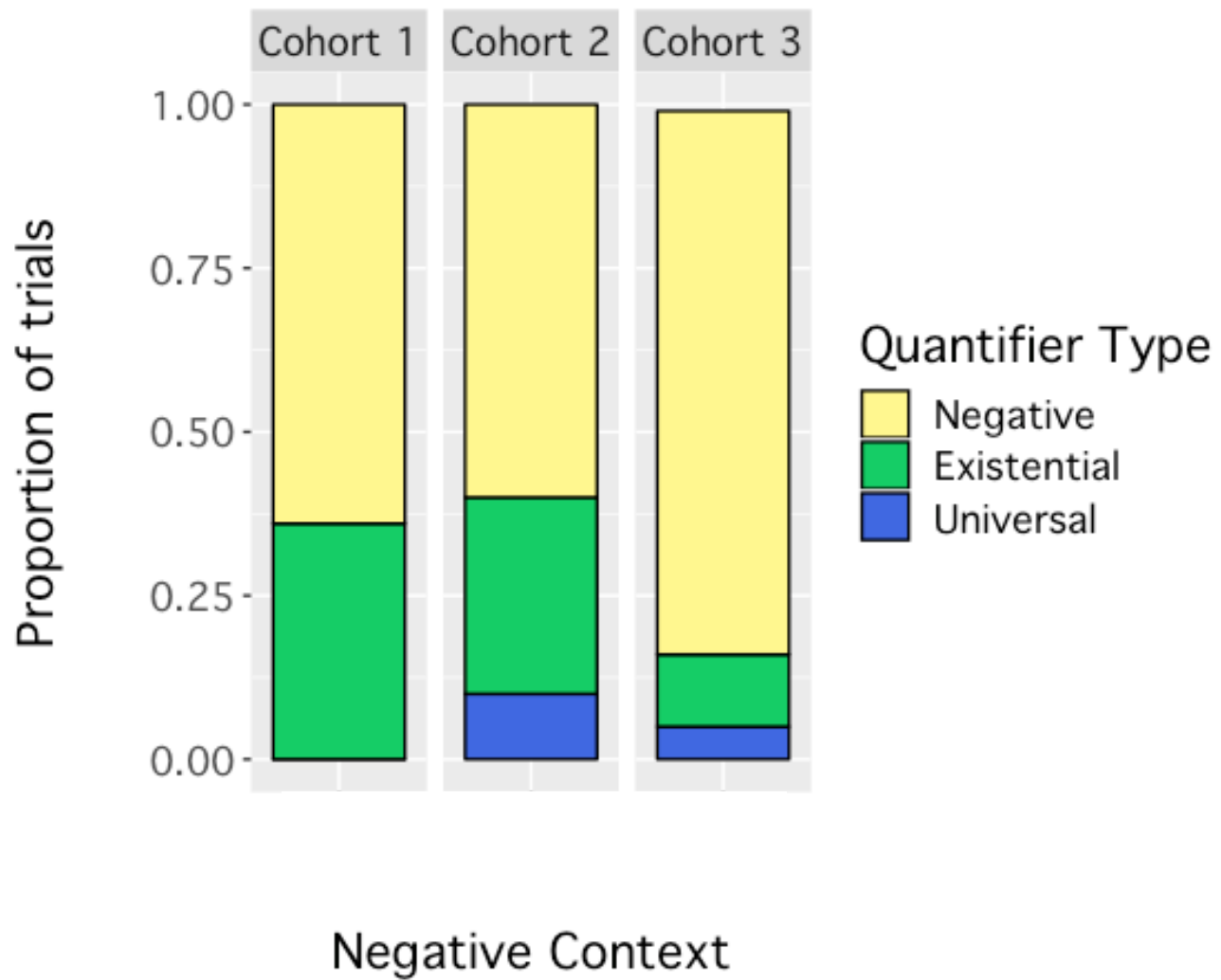
Given any X , property P is true.
(all)

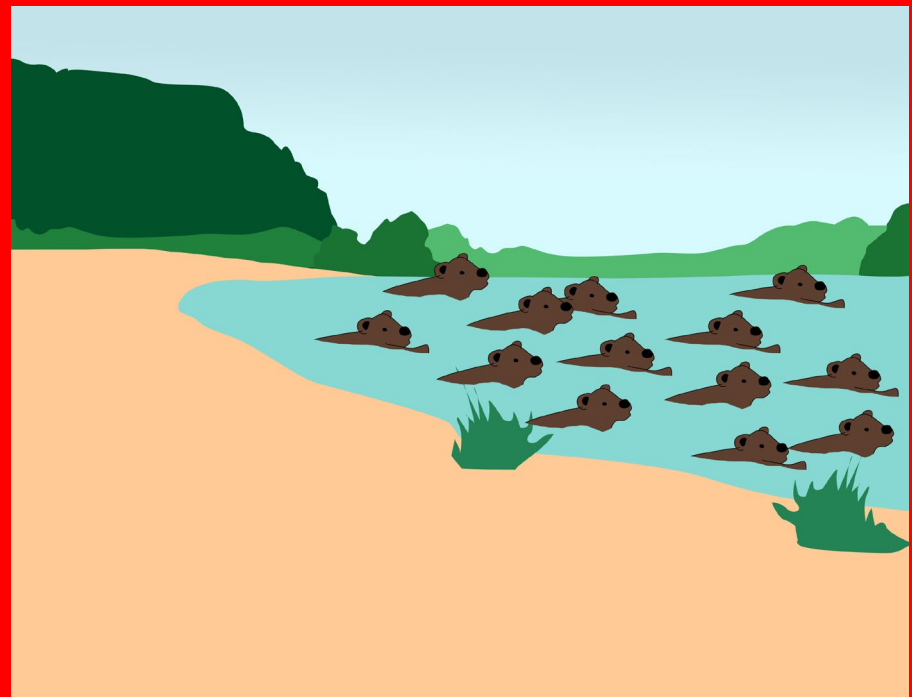
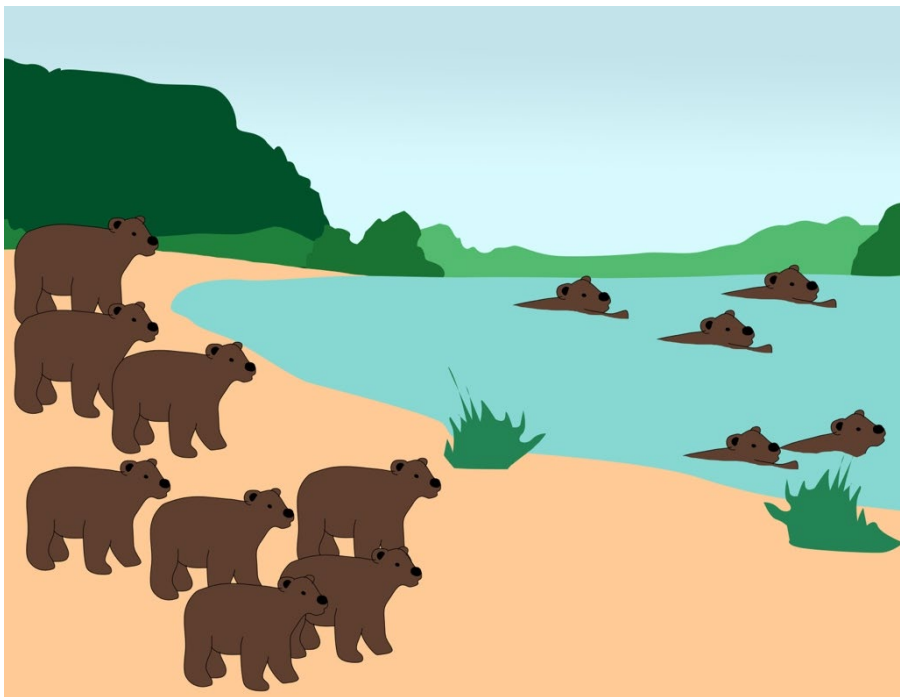
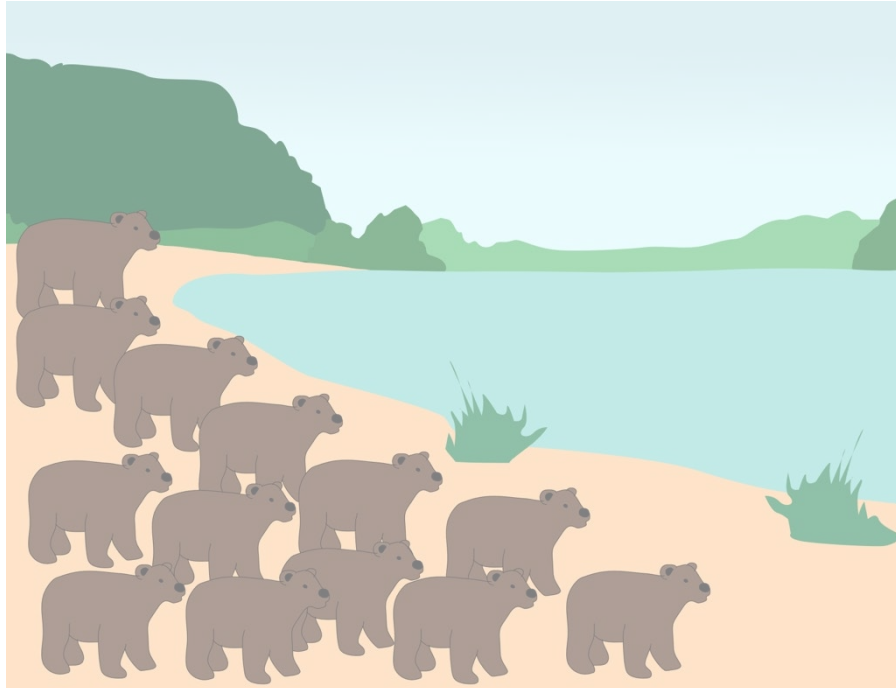


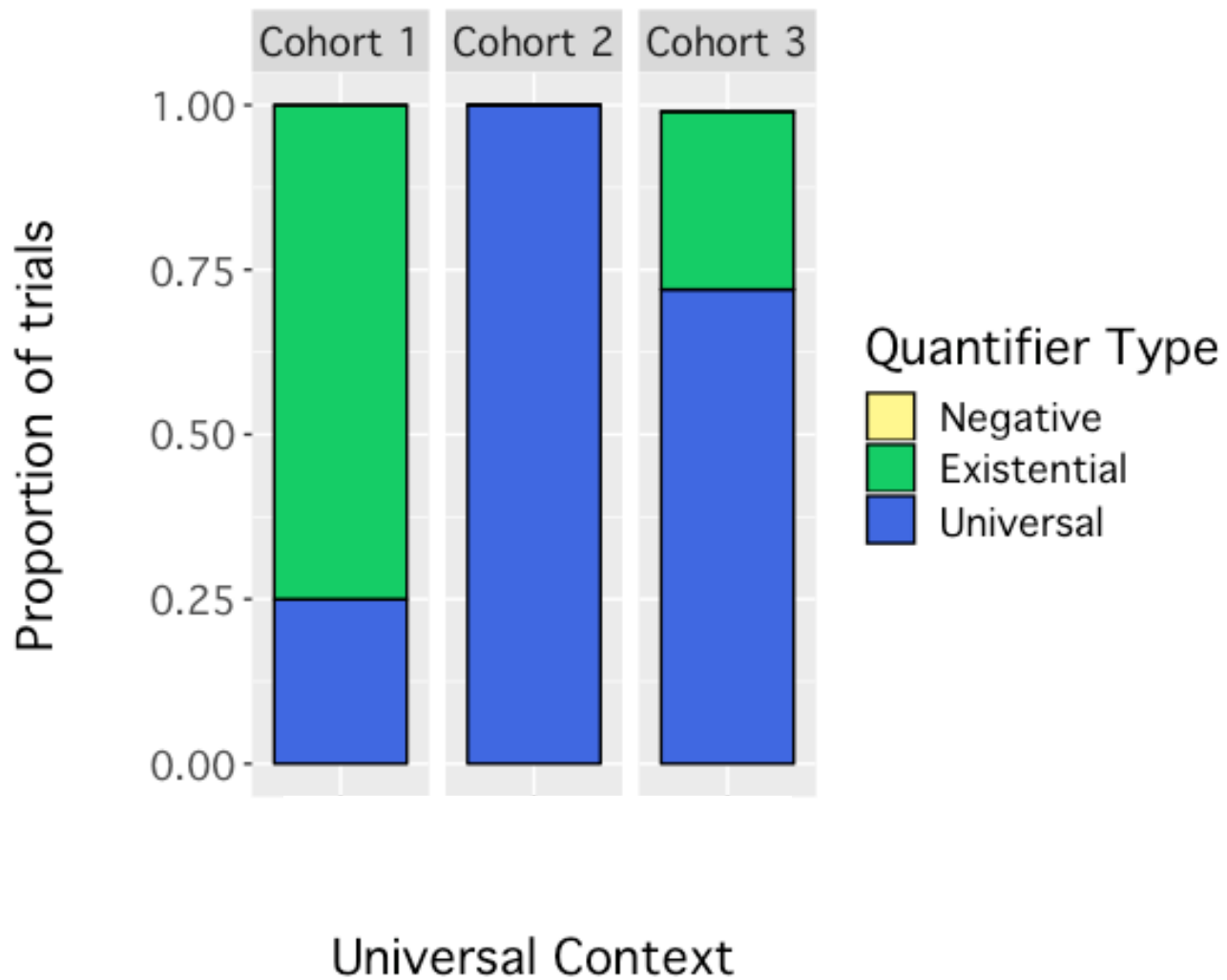












Quantifiers Conclusions

- Quantifiers emerge early in NSL
 - 27 of 28 produced an instance of each type
- First cohort speakers use existential quantifiers in universal contexts
 - Literally true but underinformative
 - Failure to consider alternatives? Less complex form?
- Unanswered questions
 - quantification in homesign (coming soon)
 - syntax of quantification

B. Recursion



Annemarie Kocab

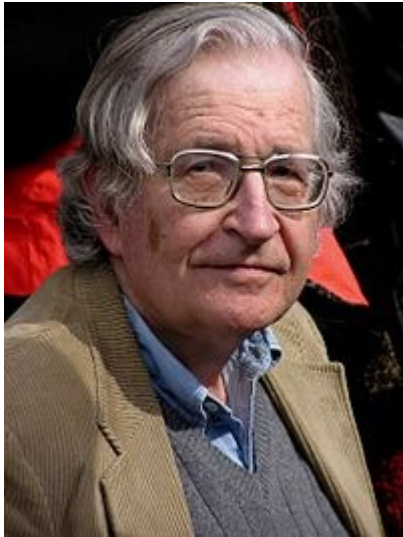


Marie Coppola



Ann Senghas

- Language makes infinite use of finite means
- Recursion: A rule that can apply to its product
- Is recursion universal?



Chomsky: Recursion is the core property of language and universal



Everett: No, recursion is absent in Pirahã



Parataxis

John has a sister. That sister has a house.

Where does recursion come from?

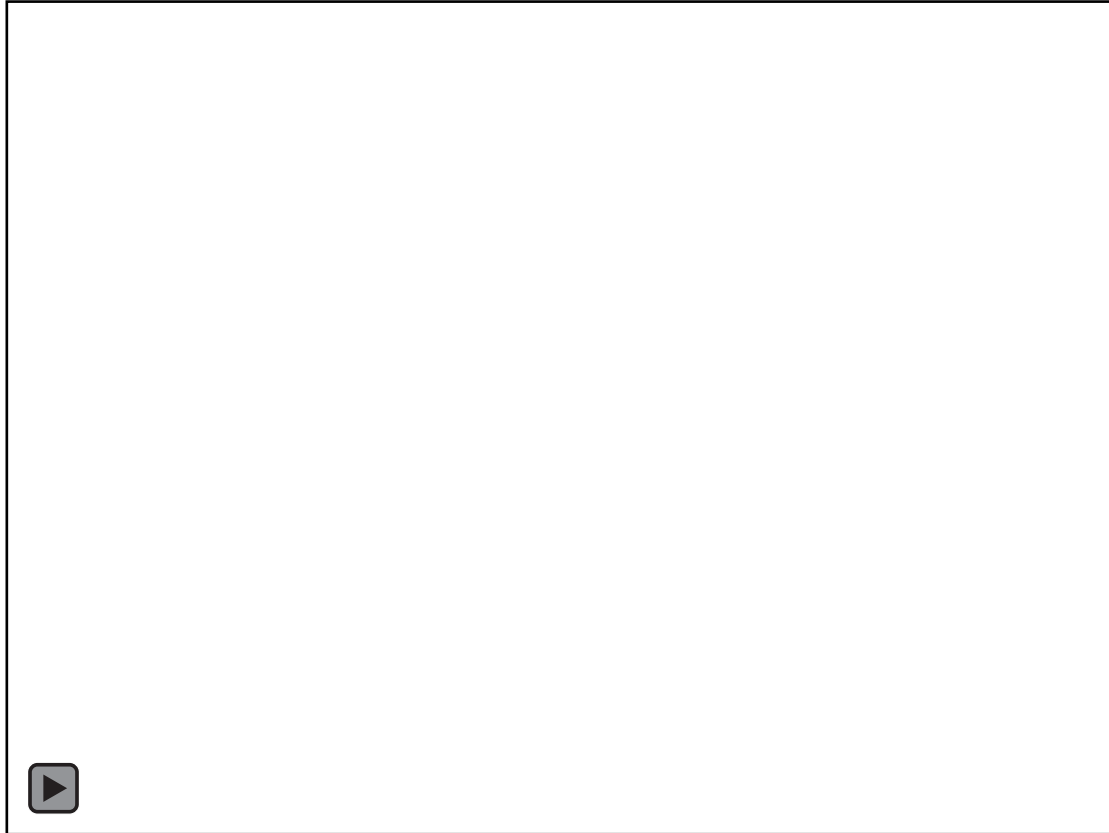
- An individual human mind, independent of linguistic experience
 - One generation
- Learning procedures that children apply in decoding a communicative system
 - A few generations
- A cultural creation (like chemistry or wheels)
 - Hundreds of generations

The questions behind the question

Which recursive representations are present?

- Semantic recursion
 - Propositions about propositions
- Syntactic recursion
 - Sentences within sentences
- Morpho-syntactic marking of recursion
 - This is the rat that lived in the house that Jack built

Elicitation Task



“The boy that was typing fell”

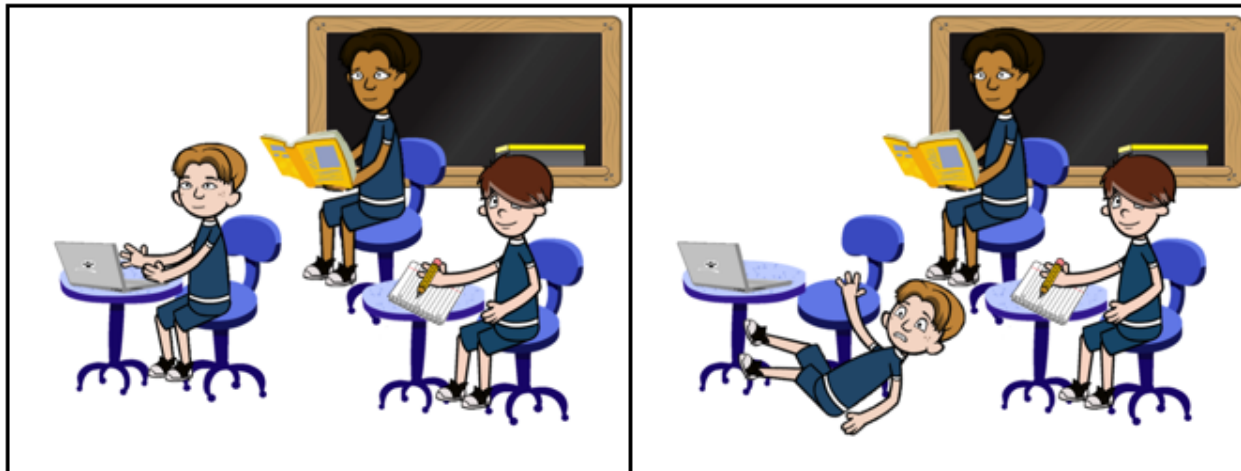
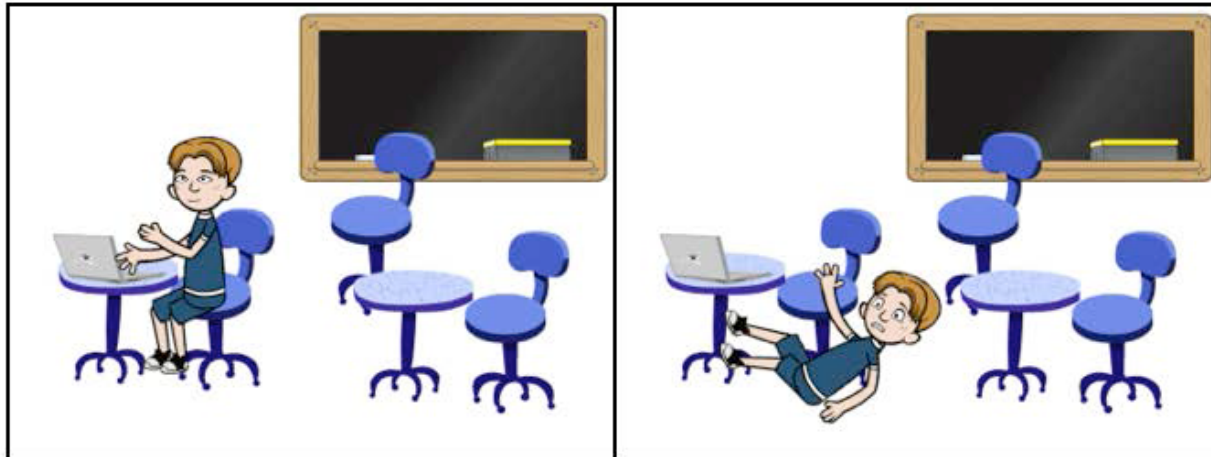
Relative clause:

There is a boy writing, a boy reading and a boy typing. **The boy (that was) typing fell.**

No Relative clause:

There is a boy writing, a boy reading and a boy typing. **He fell**

Conjoined vs. Relative



Expectations for recursive utterances in NSL

Semantics

- An utterance describing critical event with the prior event included to pick out the target individual

Morpho-syntax

- Shorter verbs in relative clauses vs. non-embedded contexts
- Visible evidence of embedding

Video deleted for privacy

Relative clause condition:

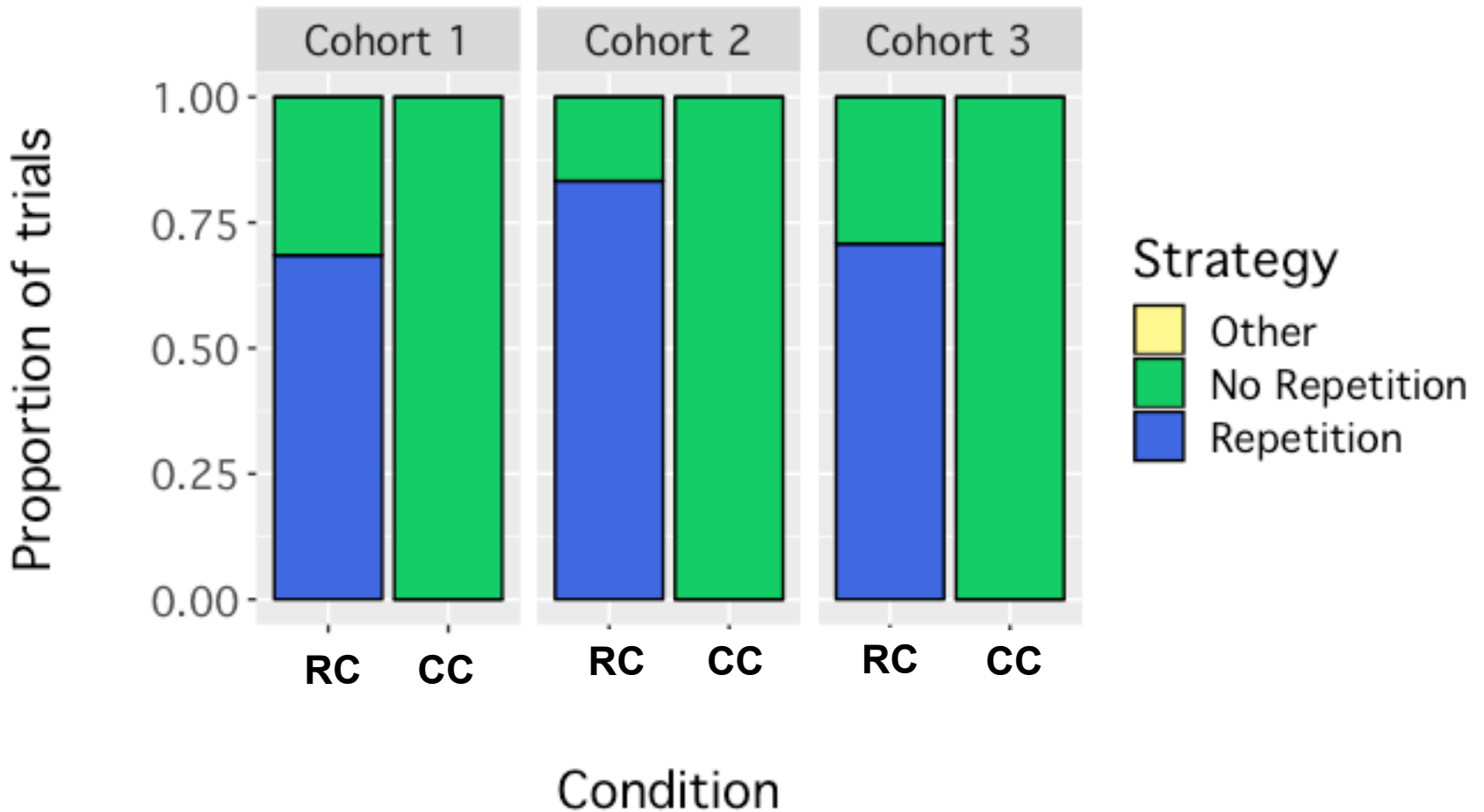
GIRL READ-BOOK, GIRL TALK-ON-PHONE, GIRL
EAT-APPLE

GIRL EAT-APPLE TAKE HAT

Conjoined clause condition:

GIRL EAT-APPLE TAKE HAT

Verbs are repeated only in the relative clause condition



Measuring relative verb length

Relative Clause



WOMAN

DRAW

REMOVE

PICTURE

Conjoined Clause



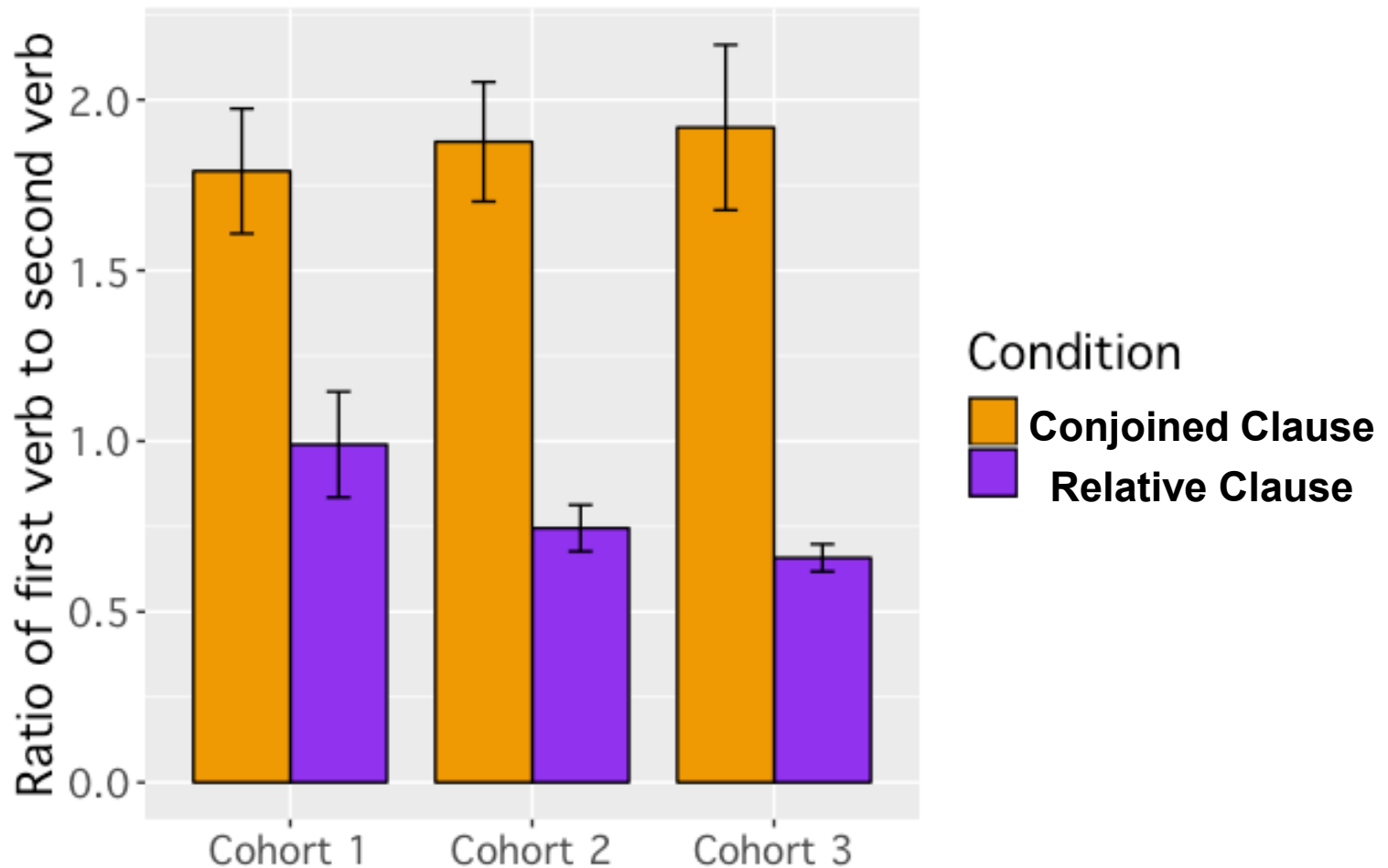
WOMAN

DRAW

REMOVE

PICTURE

Verbs are shorter when in a relative clause compared to a conjoined clause



What about homesigners?

- 4 deaf individuals, linguistically isolated
- Who use gestures to communicate with their families

Picture deleted
for privacy

Relative clause condition:

GIRL EAT-APPLE TAKE HAT

Conjoined clause condition:

GIRL EAT-APPLE TAKE HAT

No repetition of the verb, no reduction

Conclusions

- Semantic and syntactic recursion emerges in the 1st generation of NSL signers
- Marked through verb reduction
- Lack of clear recursion in homesigners
 - Like Piraha?
 - Possible parataxis: non-recursive structure conveying a recursive message

C. Event Roles



Annemarie Kocab



Irene Canudas Grabolosa



Marie Coppola



Hanna Shine



Jessica Moore

Language is a way to encode our thoughts....



The girl hit the ball.



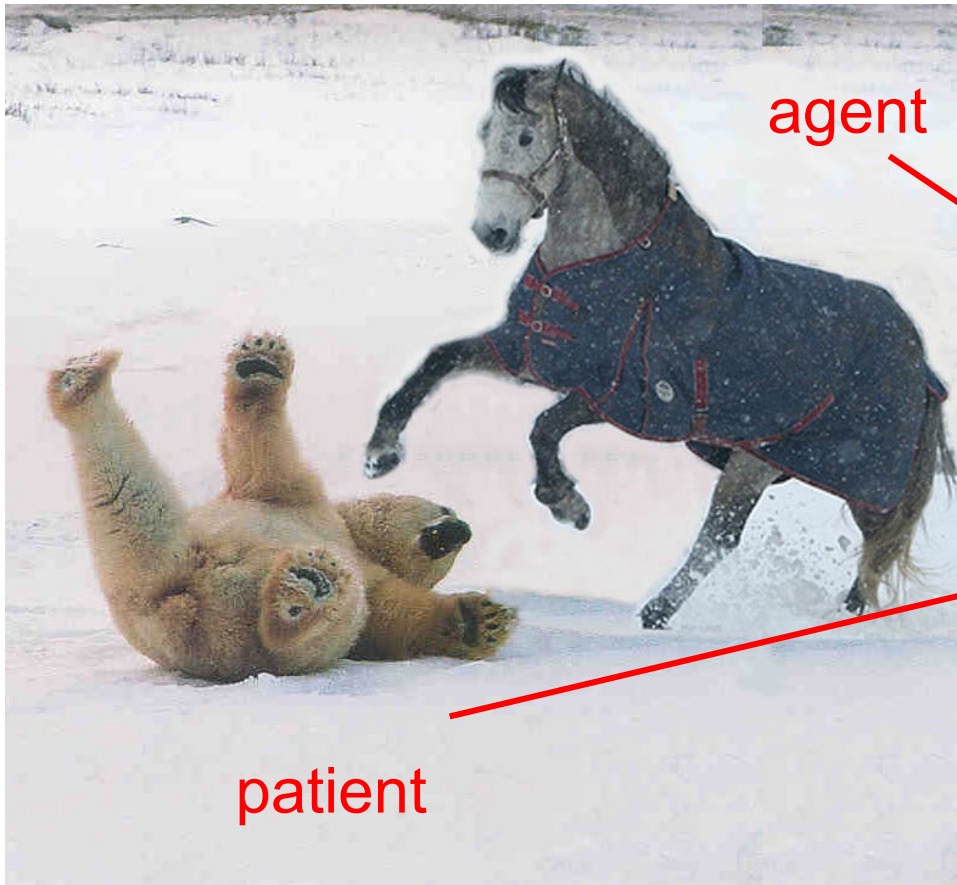


The storm trooper kicked the man



The horse *pilked* the bear.

(Baker, 1988; Marantz, 1982; Jackendoff, 1990)



The horse *pilked* the bear.

(Baker, 1988; Marantz, 1982; Jackendoff, 1990)

How widespread is the agent-patient distinction?

To determine this you need to code at the microlevel role

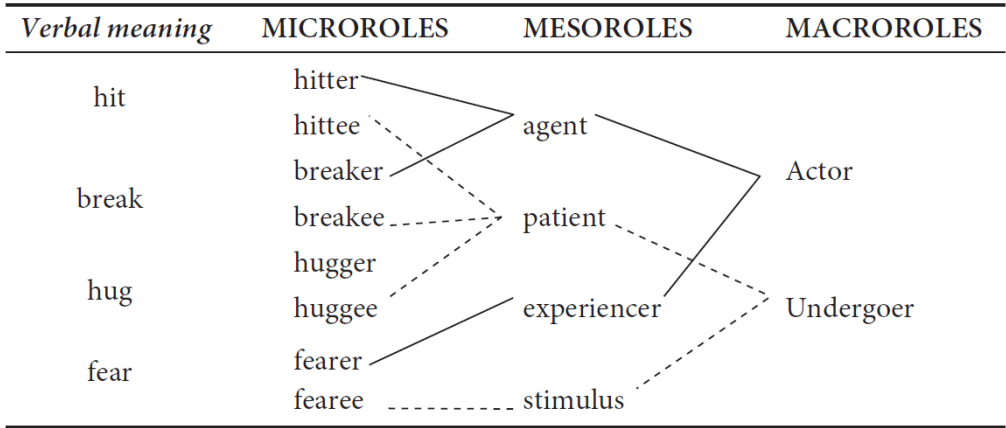


Figure 1. Hierarchy of role concepts (adapted from Van Valin 2005: 54)

Hartmann, I., Haspelmath, M., & Cysouw, M. (2014).

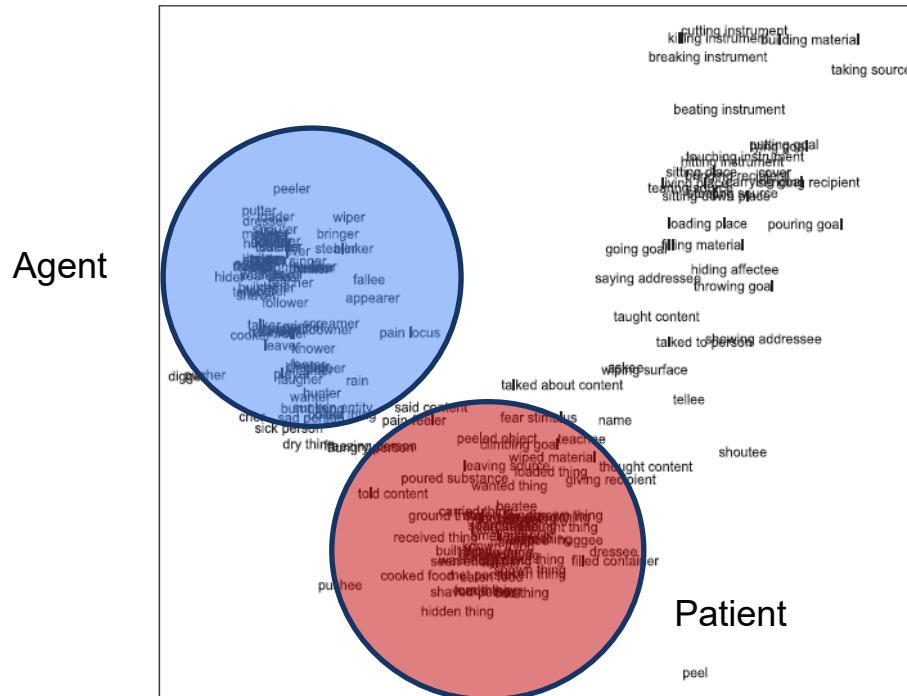
How widespread is the agent-patient distinction?

Table 1. Seven microroles with associated coding sets in three different languages⁷

microrole	Icelandic	Hoocak	Chintang
hitter	NP-nom & sbj.V	act.V	NP-erg & V.sbj
hittee	NP-acc	und.V	NP-abs & V.obj
liker	NP-dat	act.V	NP-erg & V.sbj
likee	NP-nom & sbj.V	und.V	NP-abs & V.obj
throw goal	<i>inn um</i> +NP-acc	NP+ <i>eeja</i>	NP-abs & V.obj
helper	NP-nom & sbj.V	act.V	NP-erg & V.sbj
helpee	NP-dat	und.V	NP-abs & V.obj

Hartmann, I., Haspelmath, M., & Cysouw, M. (2014).

Languages differentiate agents and patients



- 25 Languages
- 87 micro-roles
- Code the overlap between each pair of roles
- To create a multidimensional space

Figure 3. First two dimensions of a multidimensional scaling of the weighted average metric on the 87 micro-roles, showing clear semantic separation into meso- or macro-roles

Hartmann, I., Haspelmath, M., & Cysouw, M. (2014).

Where do agent/patient concepts come from?

- Historical development
 - languages differentiate meso-roles because makes for efficient communication
 - taught to children via language
 - should be absent in persons with no language model
- Human evolution:
 - should be available to homesigners despite lacking a language model

To study this we need...

- A non-linguistic task
- That is unambiguous
- Where it is plausible that role, independent of event, is relevant
- That requires generalization
 - You could have a concept of hitter, but no concept of agent

Each Trial Consists of an Event

Video Structure
(Contact)



Character Presentation

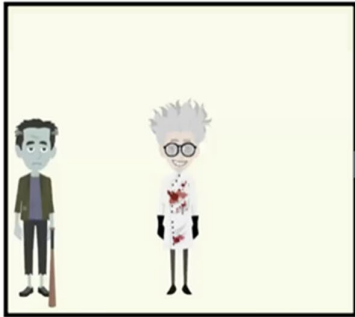


Contact Action



Character Selection

Video Structure
(Caused-Motion)



Character Presentation

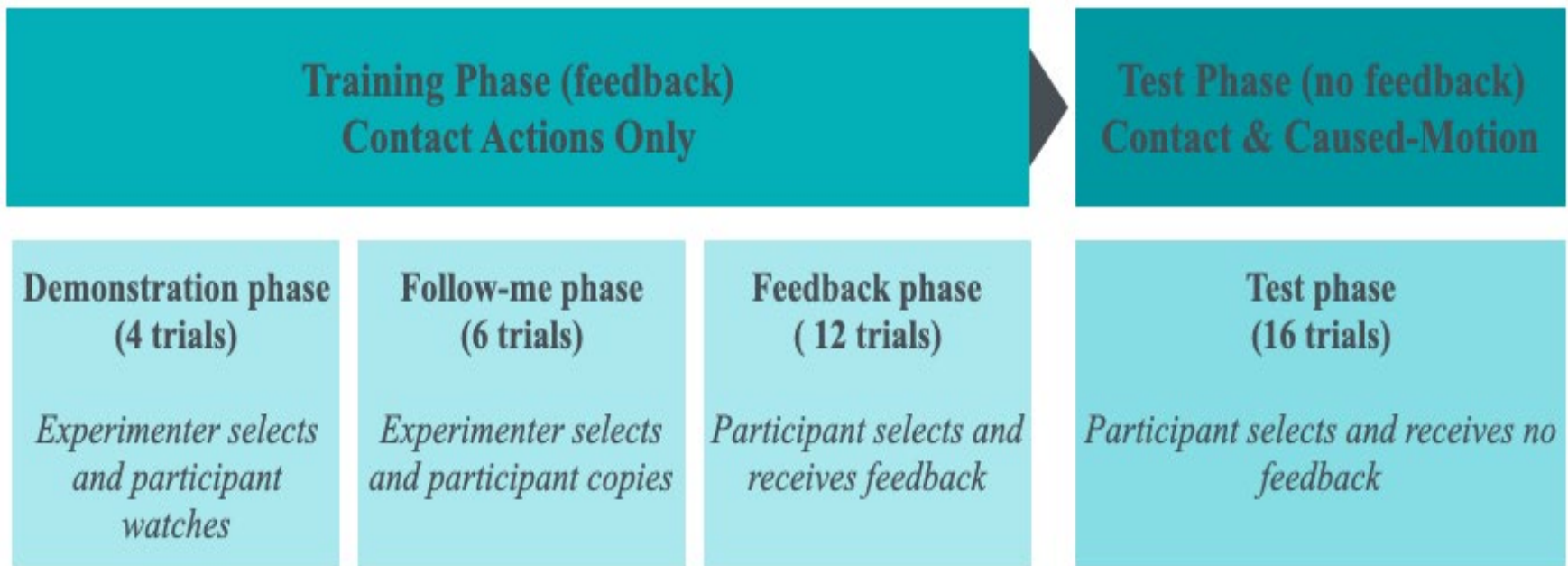


Caused-Motion Action



Character Selection

Agent-Patient: Scaffolded Imitation Task



Each Trial Consists of an Event

Video Structure
(Contact)



Character Presentation

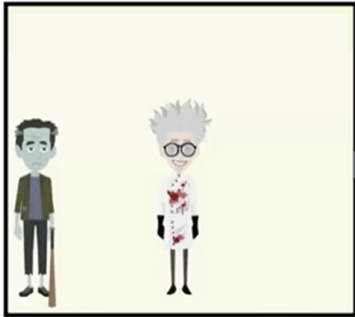


Contact Action



Character Selection

Video Structure
(Caused-Motion)



Character Presentation



Caused-Motion Action

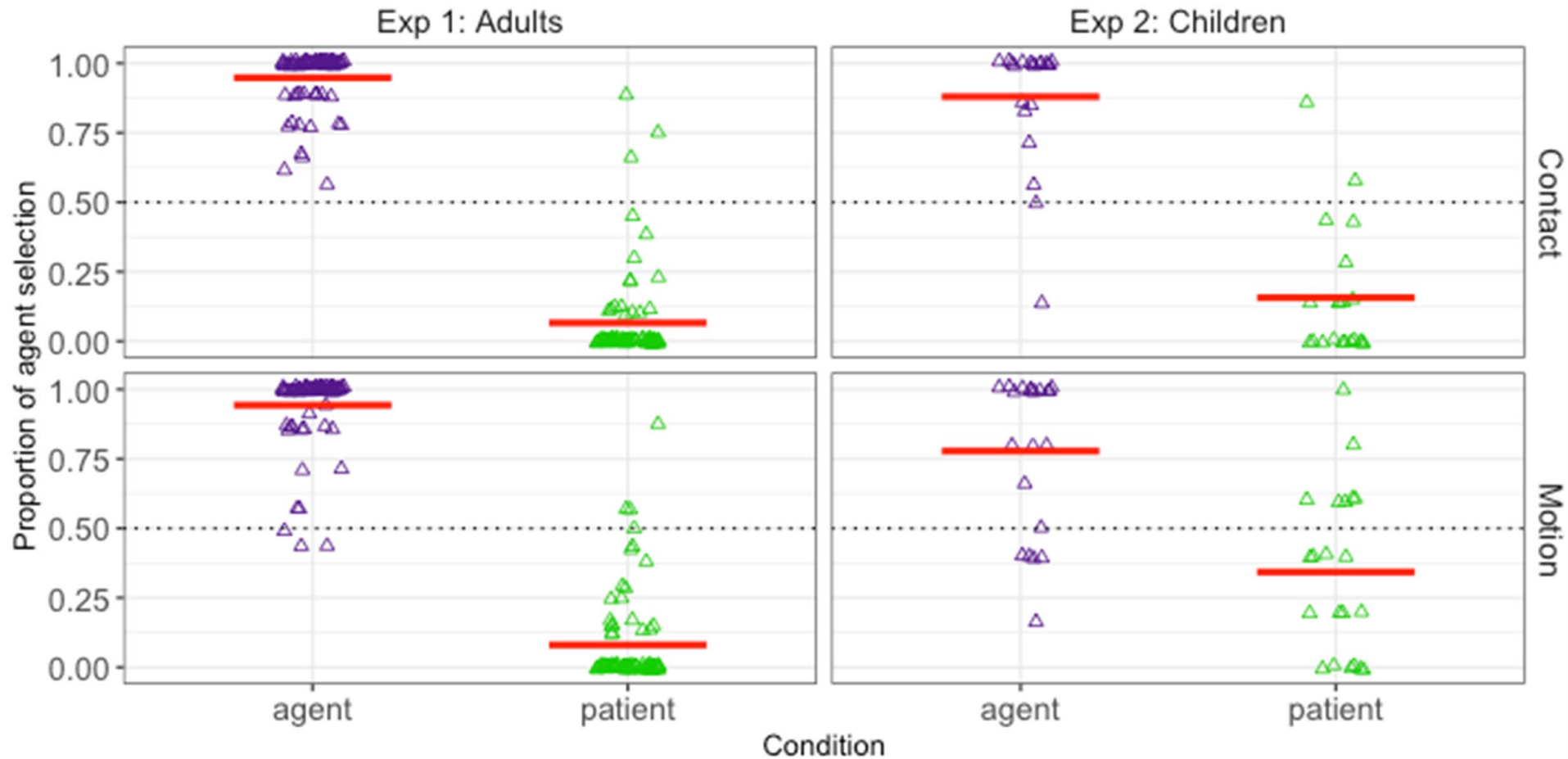


Character Selection

Our Participants

	Exp 1: Adults	Exp 2: Children	Exp 3a: Homesigners* (2024)	Exp 3b: Homesigners* (2025) <i>slower stimuli</i>
<i>N</i> <i>(after exclusions)</i>	90	20	7	7
<i>Excluded</i> <i>(failed training)</i>	75 (attention check)	5	3 (in patient condition)	No exclusions
<i>Demographics</i>	M _{age} = 34 45 women	M _{age} = 4;11 12 girls	M _{age} = 46 4 women	M _{age} = 45 3 women

Experiments 1 and 2: Adults and Children



Experiments 3a and 3b: Homesigners



Conclusions

- Our non-linguistic task systematically evokes agent and patient concepts in adults & children.
- Homesigners categorize agents and patients and generalize these concepts to new actions and a novel action class.
- **Agent-patient distinction originates in the learner**

Future Directions:

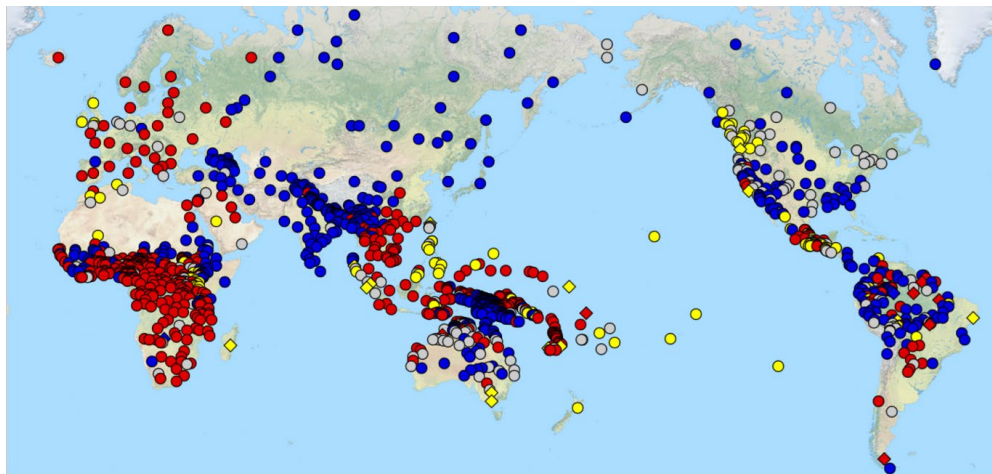
- How broad are these concepts? (More abstract event classes)
- What about other thematic roles? (Psych verbs)
- Infant studies

Are there innate syntactic
patterns that encode agentivity?

How languages encode **agents** and **patients**

- Word order (e.g., English)
 - **Man** bites **dog**
- Case marking (e.g., Turkish)
 - **Man** **dog-i** bite
- Agreement markers
 - **Men** **dog** pl-bite-sing (e.g., Georgian)
- Many languages use two or more strategies

Strong tendency for subject-first word order



- Agent-patient-verb
- Agent-verb-patient
- Verb-patient-agent
- ◆ Verb-agent-patient
- ◆ Patient-verb-agent
- ◆ Patient-agent-verb
- No dominant word order

Where do word-order patterns come from?

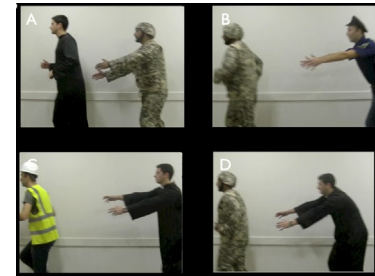
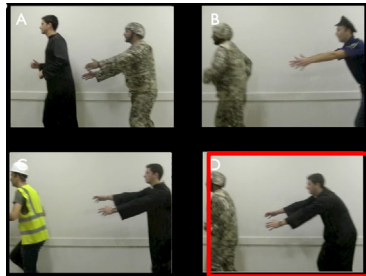
If these statistical patterns are due to syntactic evolution

- We expect Nicaraguan Sign Language and home sign
- To use word order to differentiate agents & patients
- With agent-first order

• If these patterns are due to historical processes

- We may see no systematic word order
- Variety of devices for marking argument roles

Setup of the Referential Communication Task

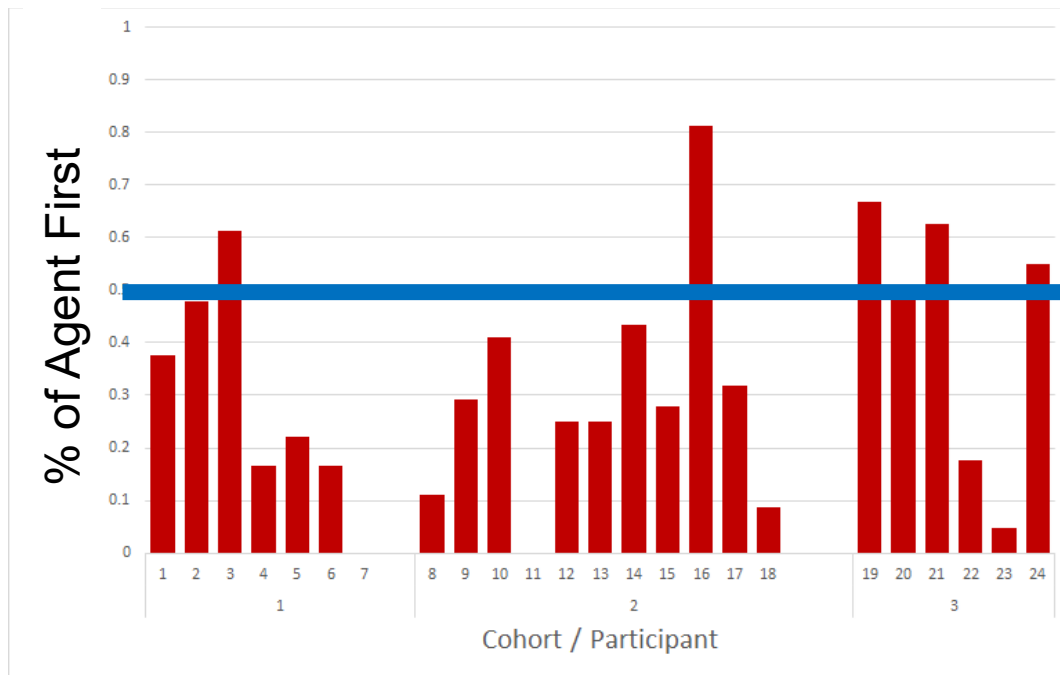


Picture deleted for privacy

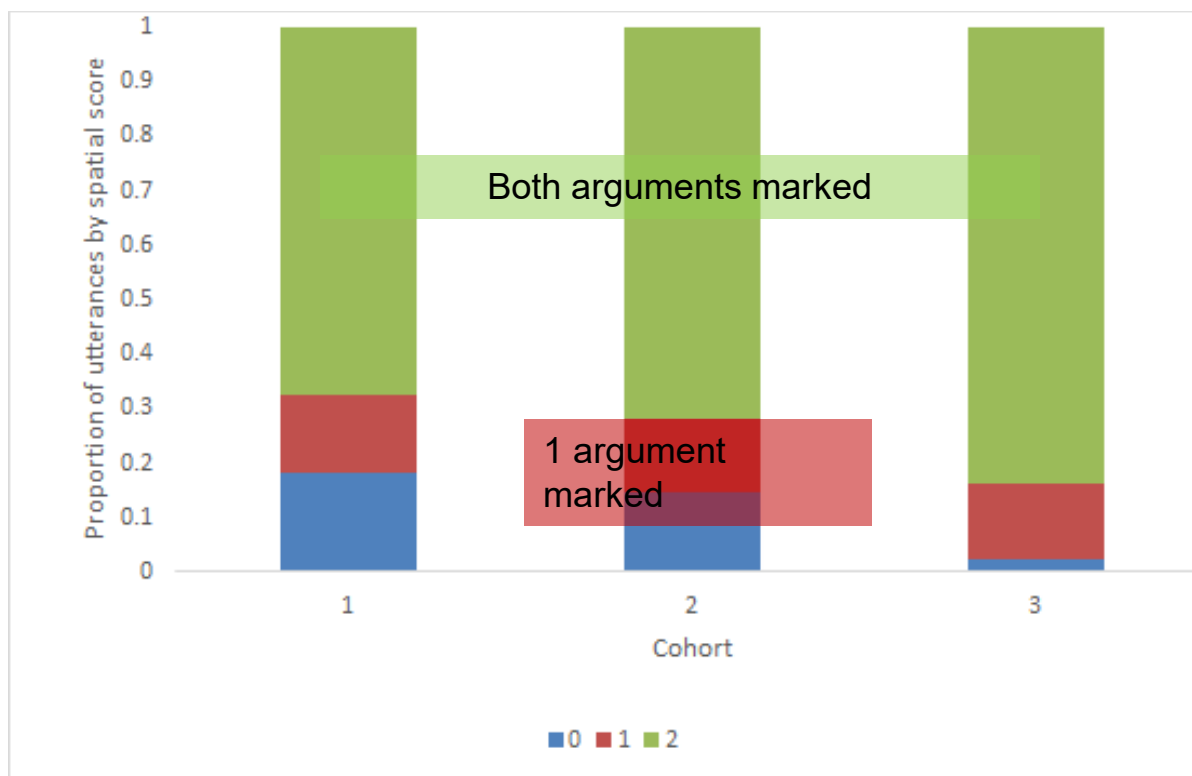
Example Display



NSL word order is highly variable



But spatial marking is widely used



Case Markers in NSL

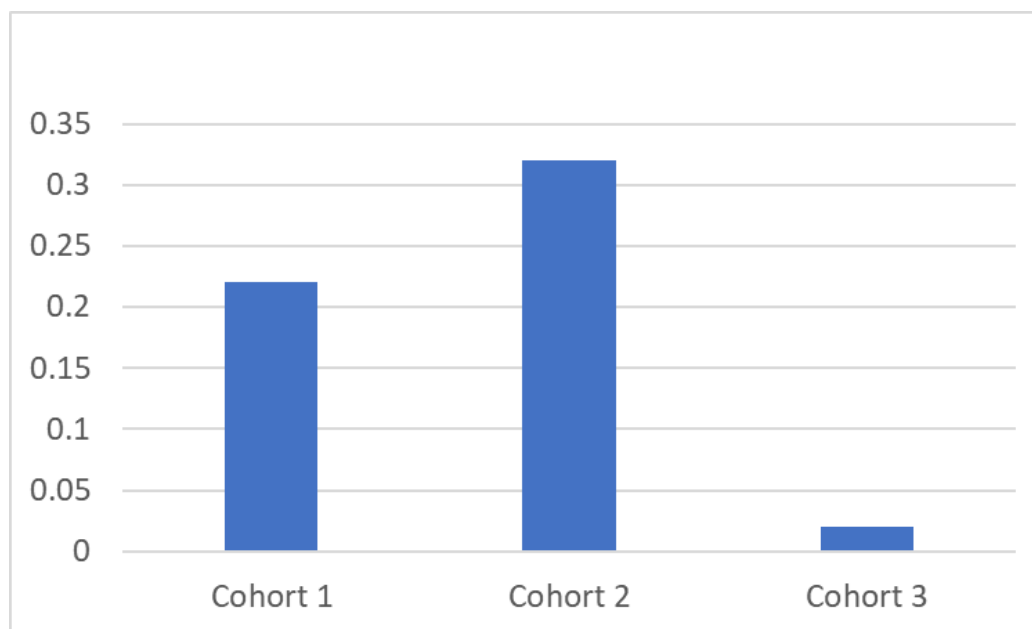
Video deleted for privacy

Cohort 1

Gloss: PRIEST marker-P FIREFIGHTER PUNCH

Event description: firefighter punch priest

Case marking disappears, as agreement becomes faster and more consistent

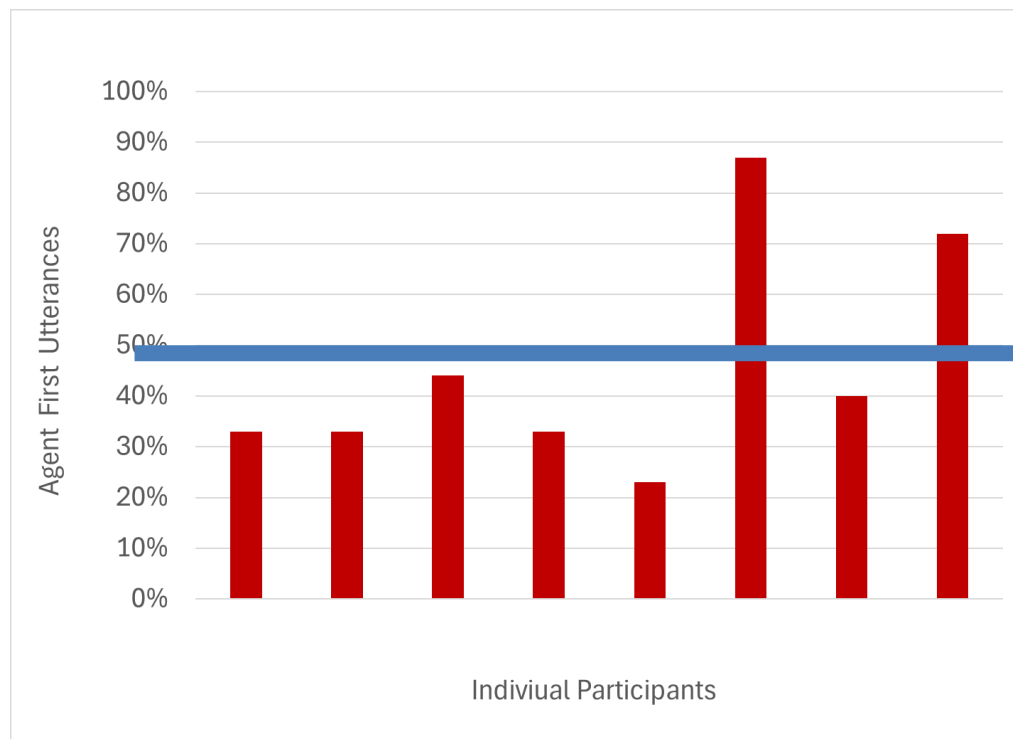


What about homesigners?

Picture deleted for privacy

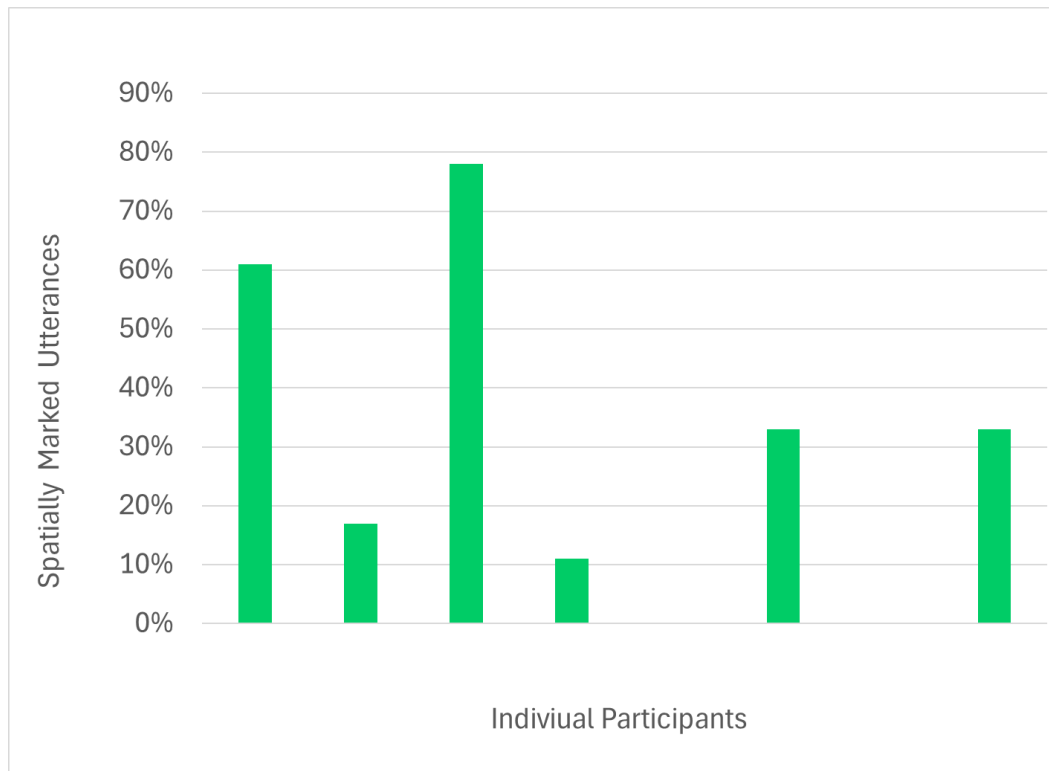
- Each homesigner is an independent case of language emergence
- The development of each system is driven by the signer

No preference for agent-first word order in homesigners

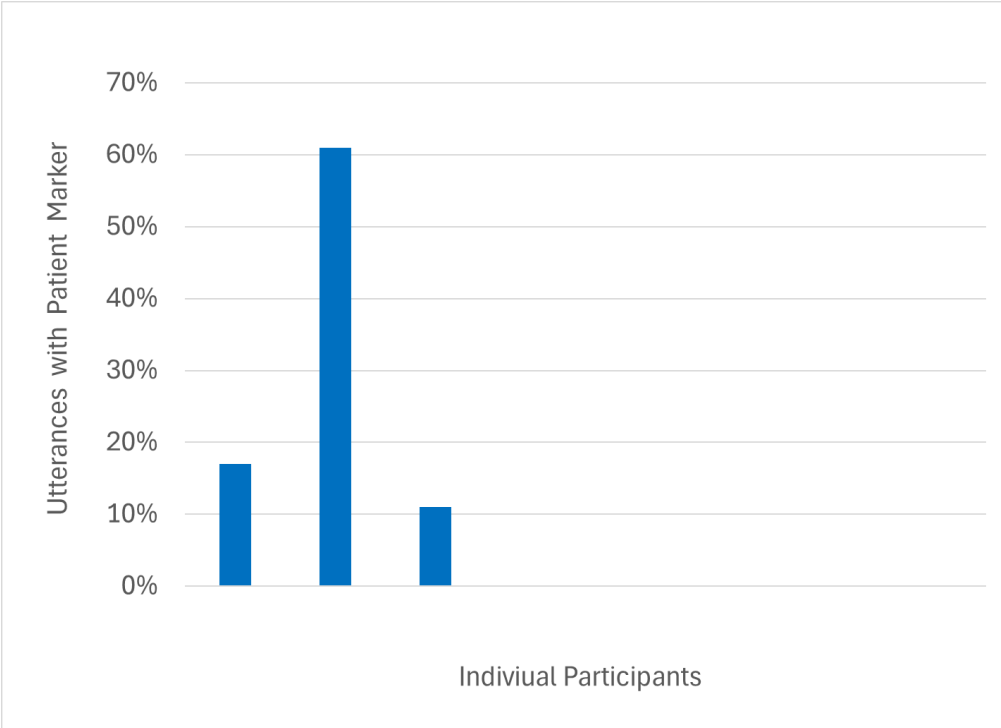


8 participants

Many homesigners use spatial marking



Three homesigners used patient markers



Event Roles Conclusions

- Do humans naturally distinguish agents and patients? **Yes**
- Do they prefer to use agent-first word order to convey this information? **No!**
- When do other devices emerge? **In 1st generation**

Roadmap

Case Studies

- Quantifiers
- Recursion
- Agents and Patients

Theoretical conclusions

- Language is shaped by evolutionary constraints
- That may be largely conceptual in nature

What are the internal forces shaping language creation?

Conceptual/Semantic

- Quantifiers
 - **Concepts are expressed in 1st generation NSL**
 - Data for homesigners forthcoming
- Recursion
 - **Expressed by first generation NSL**
 - Possibly by homesigners
- Agent/patient event roles
 - **Concepts available to homesigners**
 - **Expression of agent/patient distinction in homesign and NSL**

Syntactic

- Quantifiers
 - **Syntactic expression variable across languages**
 - Unknown for NSL
- Recursion
 - **Morphosyntactic marking variable**
 - **Phrase structure consistent**
- Agent/patient event roles
 - **No agent-first bias**
 - **Devices used vary across signers**
 - **Showing gradual shift toward spatial morphology**

What are the internal forces shaping language creation?

Conceptual/Semantic

- Quantifiers
 - Concepts are expressed in 1st generation NSL
 - Data for homesigns
- Recursion
 - Expressed by f
 - Possibly by hor
- Agent/patient event roles
 - Concepts available to n
 - Expression of agent/patient distinction in homesign and NSL

Syntactic

- Quantifiers
 - Syntactic expression variable across
- marking variable
- consistent
- agent roles
- first bias
- Devices used vary across signers
- Showing gradual shift toward spatial morphology

Language is shaped by evolved systems for conceptualization

This is not consistent with either of the dominant theories of language development!

- Constructionist Theories (Verb Island Hypothesis)
 - Infants begin by making narrow generalizations based on individual words
 - Gradually discover the semantic and syntactic categories of their language
 - Predicts that homesign and NSL should have local lexical patterns but no broad semantically guided patterns
- Linguistic Nativism (Semantic Bootstrapping, P&P)
 - Infants are guided by innate syntactic categories, underspecified rules, and innate syntax/semantics mappings
 - Predicts that homesign and early NSL should show systematic syntactic regularities consistent with innate syntax

Conceptual Nativism

If we get here....



Joshua Hartshorne
MGH Communication
Sciences & Disorders

Desiderata for a theory of acquisition

- Must account for early abstraction
- Cannot depend on extensive innate syntax

Chomsky, 1995; Dryer, 1997; Evans & Levinson, 2009; Haspelmath, 2007, 2009; Lazard, 1992

- Can exploit the statistical learning abilities of infants

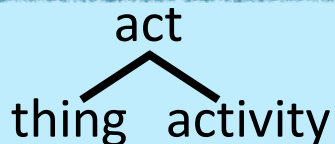
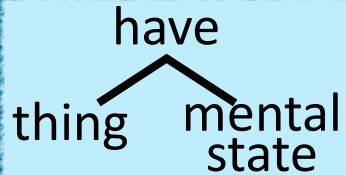
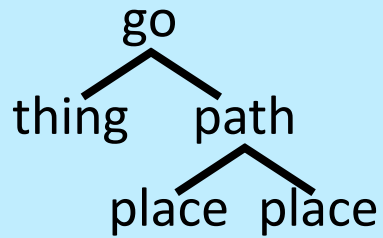
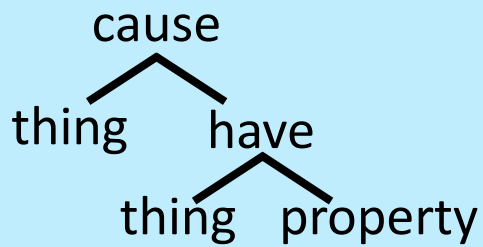
Chemla, Mintz, Bernal & Christophe, 2009; Gomez & Gerken, 2000; Gweon, Tenenbaum, & Schulz, 2010; Marcus, Vijayan, Rao, & Vishton, 1999; Marquis & Shi, 2012; Mintz, 2012; Saffran, Aslin & Newport, 1996; Saffran & Wilson, 2003; Shi & Melancon, 2010; Swingley, 2005; van Heughten & Shi, 2010

- Can exploit rich conceptual system of pre-linguistic infants (semantics)

Carey, 2009; Gergely & Csibra, 2003; Huntley-Fenner, Carey, & Solimando, 2002; Johnson, Slaughter, & Carey, 1998; Muenener & Carey, 2010; Hamlin, Wynn, & Bloom, 2010; Kuhlmeier, Bloom, & Wynn, 2004; Spelke, 1990; Spelke & Kinzler, 2007; Spelke, Phillips, & Woodward, 1995; Woodward, 1999; Leslie & Keeble, 1987.

Proposed Starting State

Semantic Structure



etc...

Clean Mapping Principle

syntactic structure reflects semantic structure

Syntactic Structure

Categories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

etc...

Structure

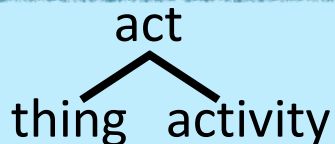
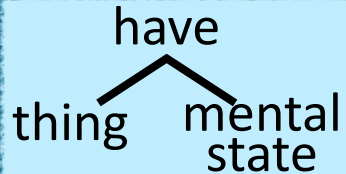
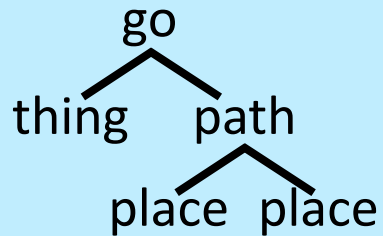
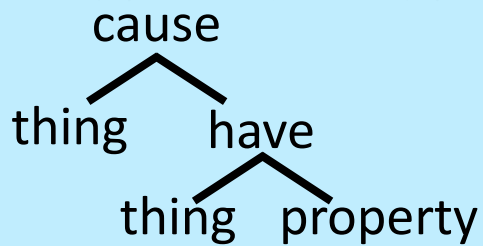
[Cluster 1 + Cluster 2]

[Cluster 3 + Cluster 4]

etc...

Infant's Starting State

Semantic Structure



etc...

Pre-linguistic Conceptual Structures

Compositional
Hierarchical
Representations

Relevant to nonlinguistic analysis of
intentional action and events

Infant's Starting State

Pattern Finding Algorithms

Domain-General

Distribution over lexical types: Find candidate categories

Distributional over categories: Find candidate rules

Syntactic Structure

Categories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

etc...

Structure

[Cluster 1 + Cluster 2]

[Cluster 3 + Cluster 4]

etc...

Infant's Starting State

Evidence

Corpus Analyses & Modelling: Cartwright & Brent, 1997; Redington, Chater & Finch, 1998; Mintz, 2003; Swingley, 2005; Chemla, Mintz, Bernal & Christophe, 2009; Connor, Fisher & Roth, 2014

Infant Studies:

Gómez & Gerken, 1999; Gómez, 2002; Saffran & Wilson, 2003; Gómez and Maye, 2005; Mintz, 2006; Shi & Melancon, 2010; van Heughten & Shi, 2010; Cyr & Shi, 2012; Mintz, 2012

Syntactic Structure

Categories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

etc...

Structure

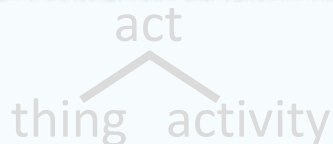
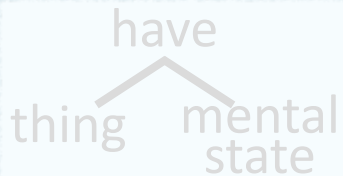
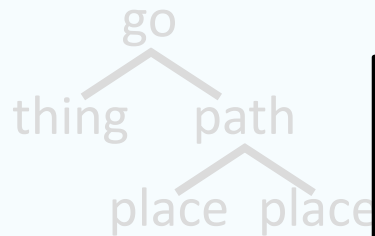
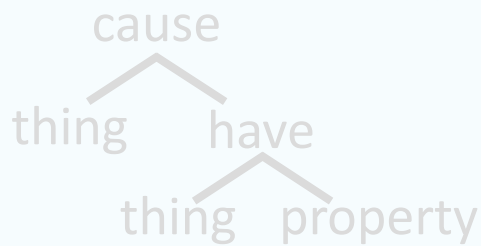
[Cluster 1 + Cluster 2]

[Cluster 3 + Cluster 4]

etc...

Infant's Starting State

Semantic Structure



etc...

Clean Mapping Principle

Assume
syntactic structure reflects
semantic structure

An idea with a long history:

Montague, 1970; Baker, 1988; Pinker, 1984; 1988;
Gleitman, 1990; Fisher, Hall, Rakowitz & Gleitman,
1994; Macnamara, 1982 inter alia

Structural Isomorphism:

Bouchard, 1995; Jackendoff, 1992; Levin &
Rappaport Hovav, 2005; Wechsler, 1995; Williams,
2003

Syntactic Structure

Categories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

etc...

Structure

[Cluster 1 + Cluster 2]

[Cluster 3 + Cluster 4]

etc...

Infant's Starting State

Semantic Structure

cause

th

t

thi

act

thing activity

etc...

Syntactic Structure

Categories

2

4

6

..

[Cluster 3 + Cluster 4]

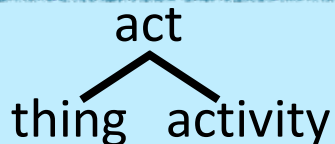
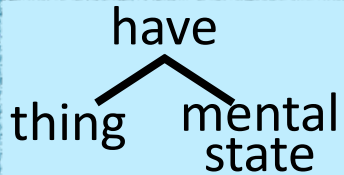
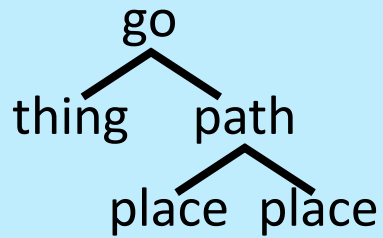
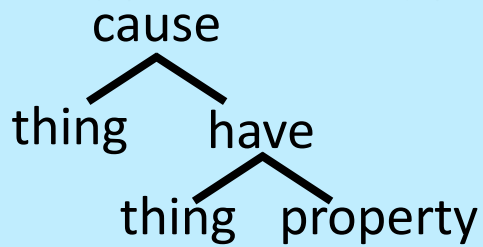
etc...

Clean Mapping Principle

- ***Domain-general mapping algorithm***
 - Category-to-category mapping
 - Structure-to-structure mapping (preserve dominance)
- ***Domain-specific expectation***
 - patterned intentional signals reflect conceptual structure

Infant's Starting State

Semantic Structure



etc...

Clean Mapping Principle

syntactic structure reflects semantic structure

Syntactic Structure

Categories

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

etc...

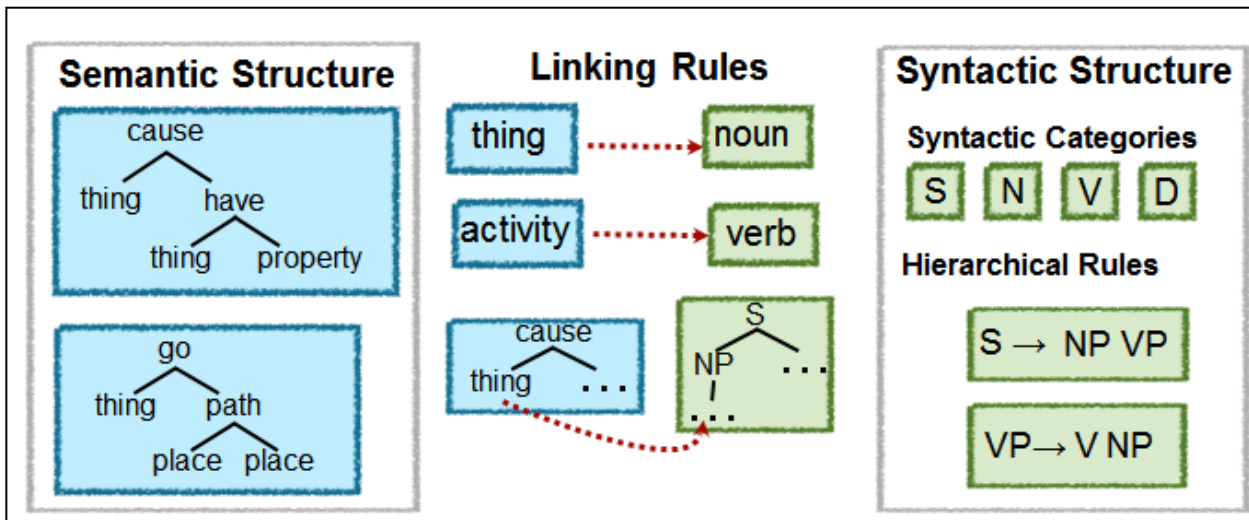
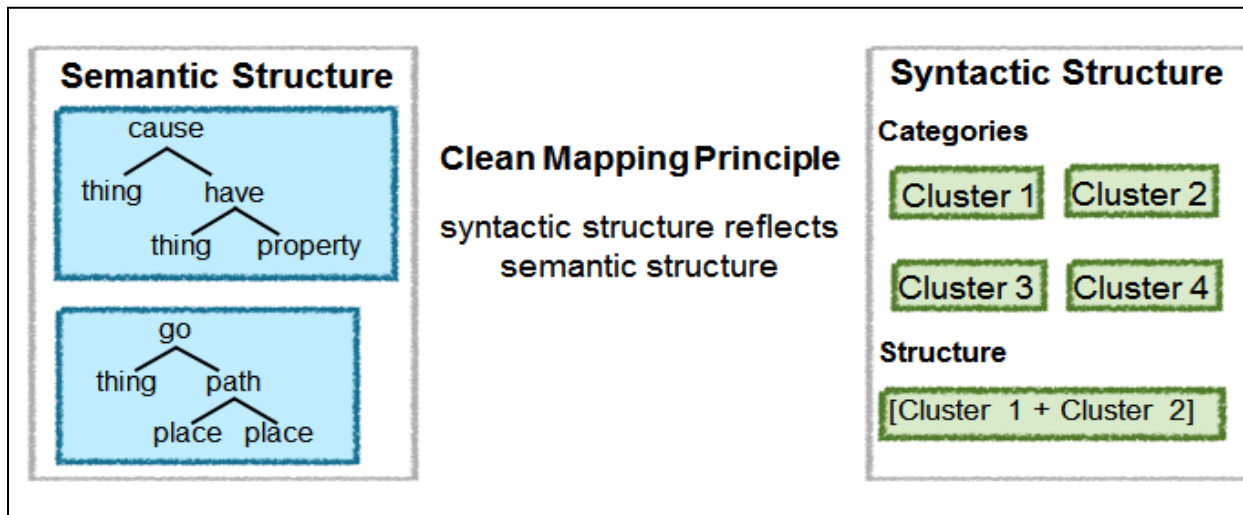
Structure

[Cluster 1 + Cluster 2]

[Cluster 3 + Cluster 4]

etc...

Developmental Story



Not Counter Evidence (Things that I fully expect)

- Syntactic universals related to meaning (nouns and verbs)
- Systematic syntactic variation due to ambiguity at the semantic/syntactic interface
 - OV or VO; ergative vs nominative
- Differences in performance of adults and children on conceptual tasks
 - Having a concept is just the beginning

Potential Counter Evidence

(Things that would nibble away at the hypothesis)

1. Syntactic universals unrelated to semantics
2. Syntactic parameters that cannot be explained by ambiguity at the semantic/syntactic interface or logical necessity
3. Substantive semantic variation across languages
4. Evidence that linguistically-isolated individuals lack concepts encoded in semantic structures

Most deaf people in the world are homesigners



**WORLD FEDERATION
OF THE DEAF**

70% of deaf people in the world lack access to language and education

~90-95% in Nicaragua

*World Federation of the Deaf, 2016;
Coppola, 2020*



Señas y Sonrisas

A smartphone app
that helps parents
learn basic NSL
signs to
communicate with
their deaf children



WHAT CAN YOU DO??

- Challenge assumptions that language is only spoken
- Remove barriers for deaf people in academia
- Raise awareness of language deprivation

We thank the signing community in Nicaragua, the homesigners and their family members for allowing us to learn from them.

Pictures deleted for privacy



This material is based upon work supported by the National Science Foundation under grant # 2116180. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.