Utrecht University

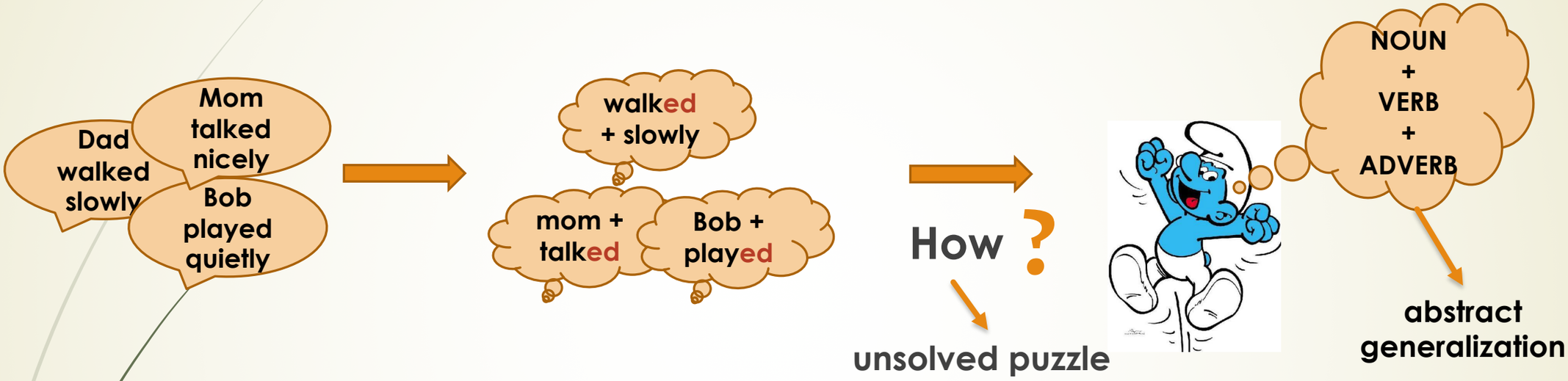
When Attention Distraction Helps Rule Induction

An Entropy Model

UiL OTS
UTRECHT INSTITUTE OF LINGUISTICS

Silvia Rădulescu & Mridhula Murali
Sergey Avrutin, Frank Wijnen
Utrecht University

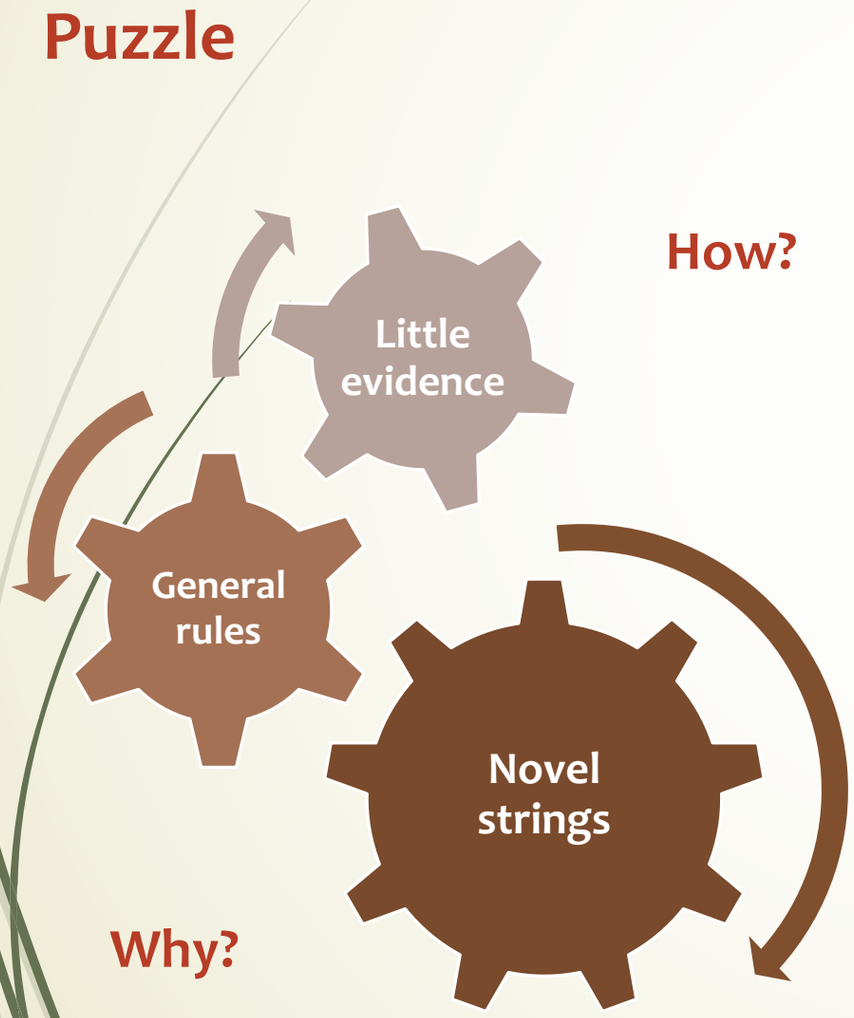
How do we make generalizations from little evidence?



? How do learners converge on abstract generalizations?

What triggers the inductive leap from memorizing to rule induction? ?

Types of Rule Induction (generalization)



Item-bound generalizations

→ relations between specific items

- e.g. *verb* + ***"-ed"***

Category-based generalizations

→ operations beyond specific items

→ over abstract categories

- e.g. **NOUN + VERB + ADVERB**

Independent underlying mechanisms ?

1. • **Statistical learning -> Item-bound generalizations**
   - *ba* follows *ba*, end in *di*
   - Saffran et al. (1996); Aslin et al. (1998)

2. • **Abstract rule learning -> Category-based generalizations**
   - *varX* follows *varX*, end in *varY*
   - Marcus et al. (1999)

OR • **Statistical learning -> BOTH item-bound & category-based generalizations**
- Aslin & Newport (2012); Frost & Monaghan (2016)

OR
Phased mechanism?

- *le_le_di*
- *de_de_je*
- *ko_ko_ba*

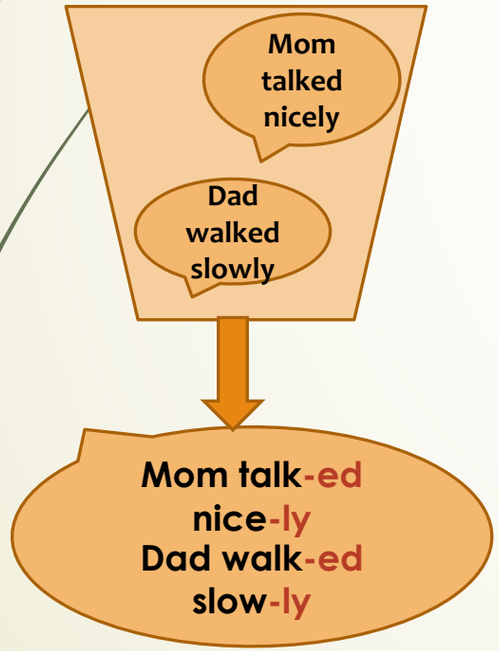
- *di* after 2x *le*
- end in je, *ba* or *di*

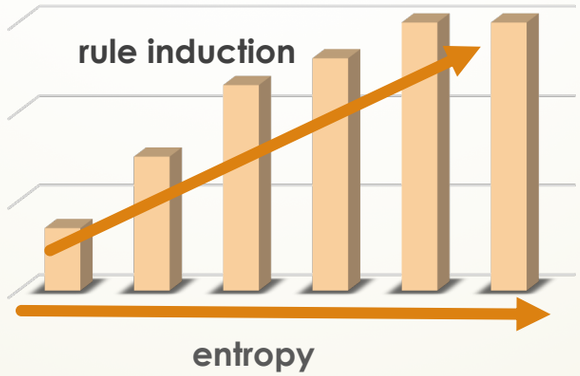
XXY

Entropy Model

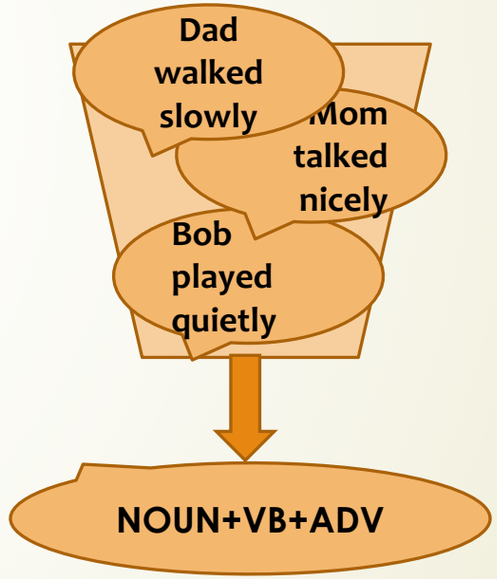
Rule Induction → the interaction between ***input complexity*** **(entropy)** and the limited encoding power of the human brain (***channel capacity***)

Low complexity (entropy) → item-bound generalizations





High complexity (entropy) → category-based generalizations



? What is the effect of Input Complexity on Rule Induction ? ?

Previous research

- input complexity (variability) plays a crucial role in rule induction √
  - Gomez (2002); Gerken (2006); Reeder, Aslin & Newport (2013)
- NOT mere variability, BUT a specific pattern of variability

Question

- How can we capture this specific pattern of variability by incorporating all variables ? ?

? What is the effect of **Input Complexity** on Rule Induction ? ?

→ vary **Input Complexity** & keep **Channel Capacity** constant

Artificial Grammar Learning - Experiment 1 + 2

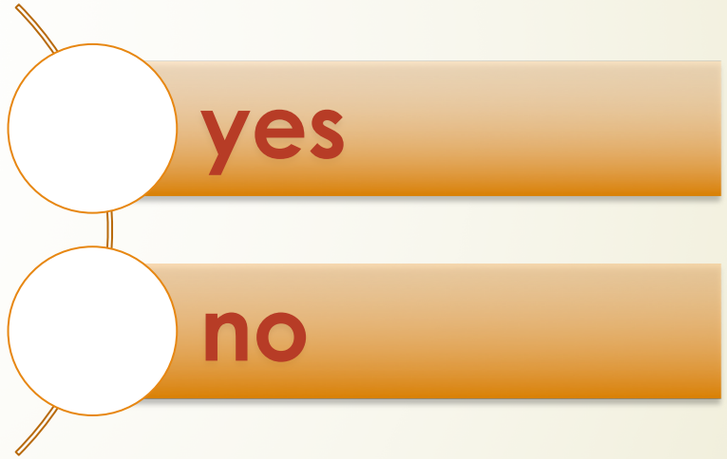
- **adults, ~22y, ~4min, between-subjects**
- **3-syllable XXY: *daa_daa_lie***
- **manipulated *ENTROPY* (number & frequency)**
  - **LowEN**
    - 2.8 bits (4 × 7Xs / 4 × 7Ys)
    - 3.5 bits (4 × 6Xs / 4 × 6Ys)
  - **MedEN**
    - 4 bits (2 × 12Xs / 2 × 12Ys)
    - 4.25 bits (2 × 14Xs / 2 × 14Ys)
  - **HiEN**
    - 4.58 bits (1 × 24Xs / 1 × 24Ys)
    - 4.8 bits (1 × 28Xs / 1 × 28Ys)

Test

Could this string be possible in the language that you heard?

- *5 x 4 types = 20 strings*

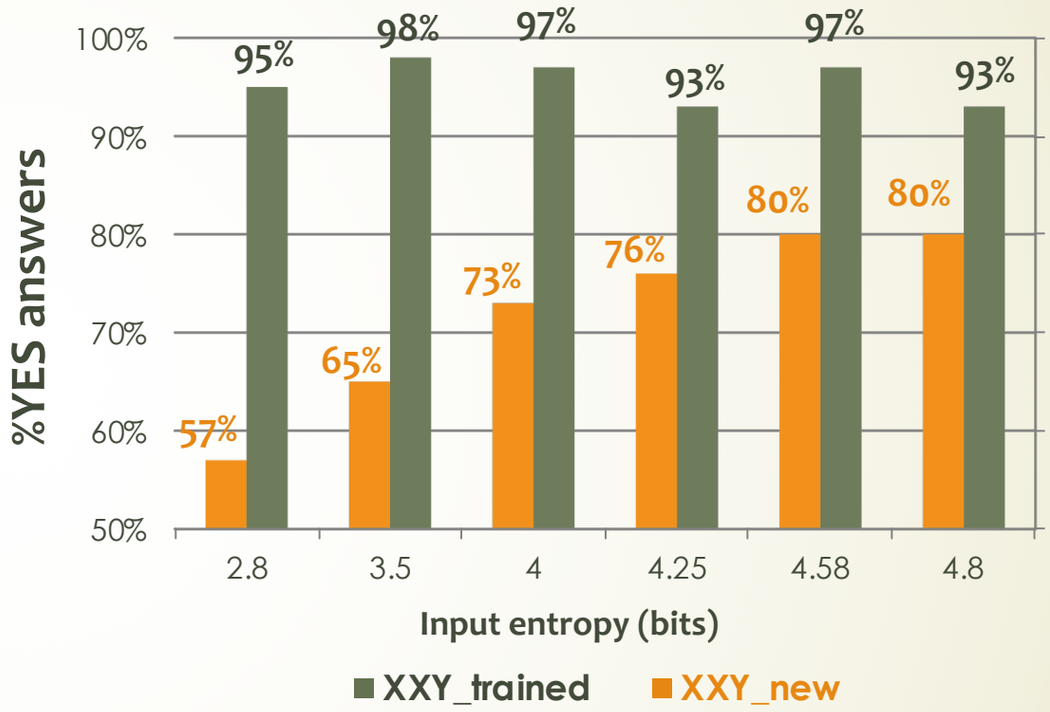
- XXY_new_syll: *too_too_suu* √
- XXY_trained_syll: *daa_daa_lie* √
- X1X2Y_trained_syll: *teu_duu_saa**
- X1X2Y_new_syll: *reu_loo_gee* *



Results

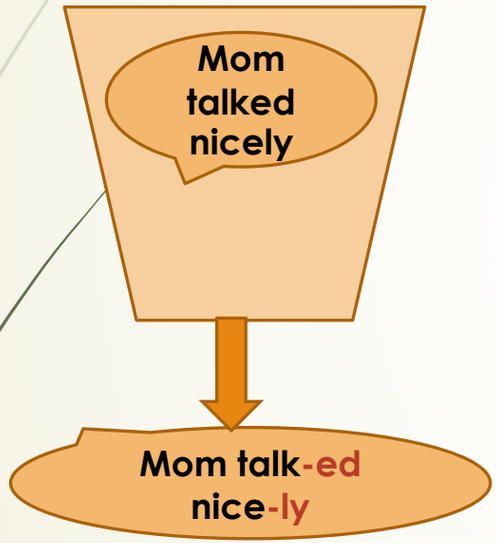
- the higher the entropy, the higher the tendency to accept **new *XXY*** strings

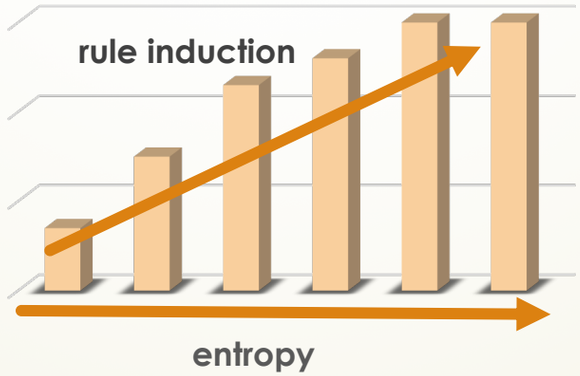
- a very similar high acceptance of ***XXY* trained** strings

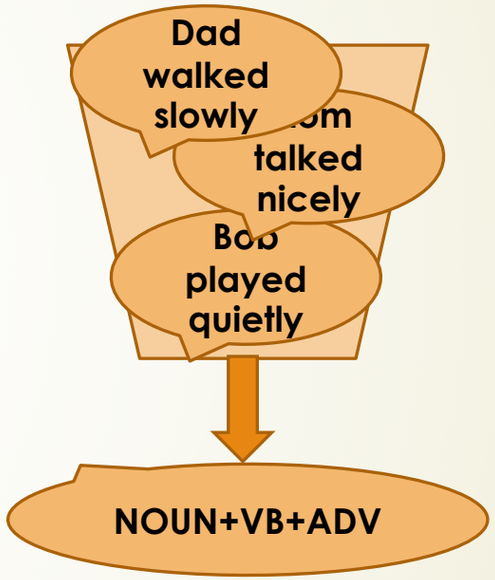


Entropy Model

Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***



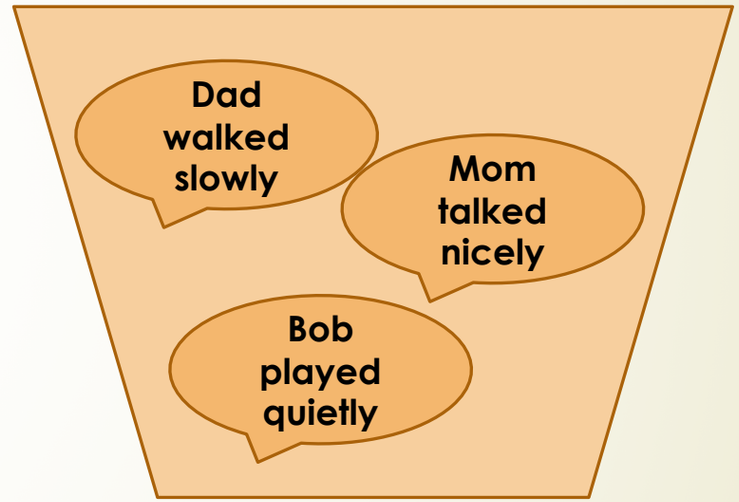


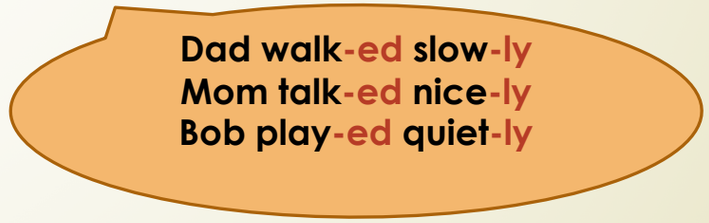


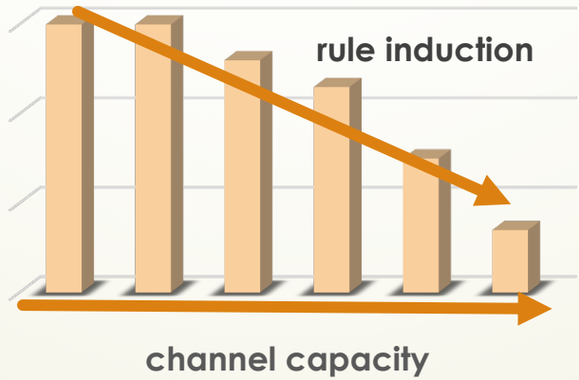
Entropy Model - hypotheses

Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

High channel capacity →
item-bound generalizations



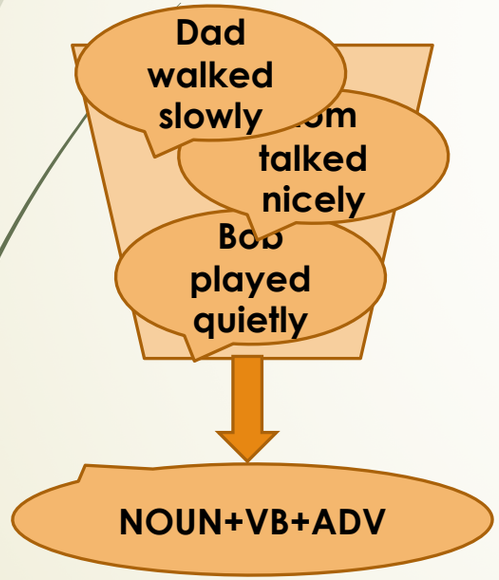


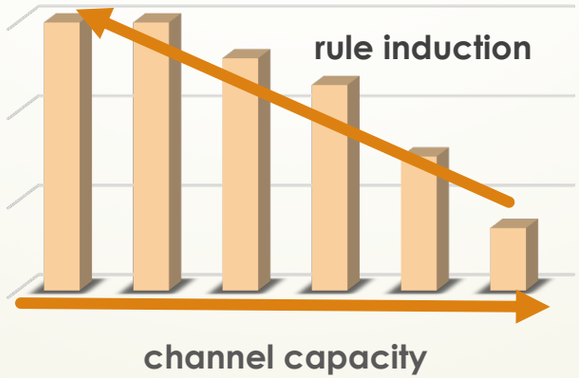


Entropy Model - hypotheses

Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

Low channel capacity →
category-based generalizations

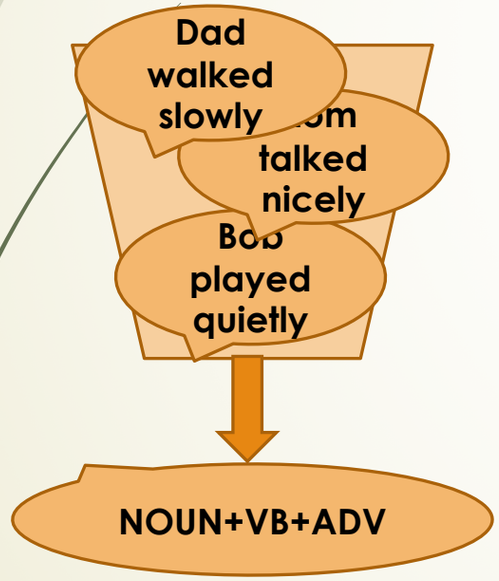


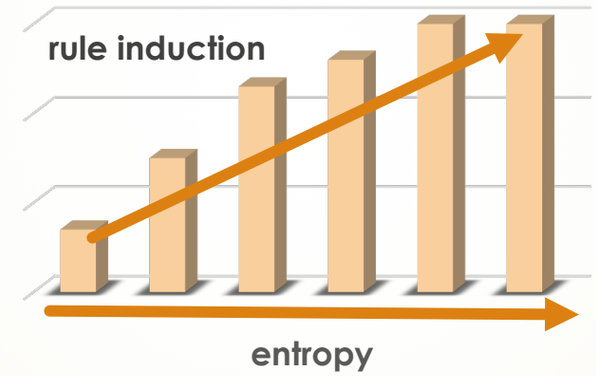


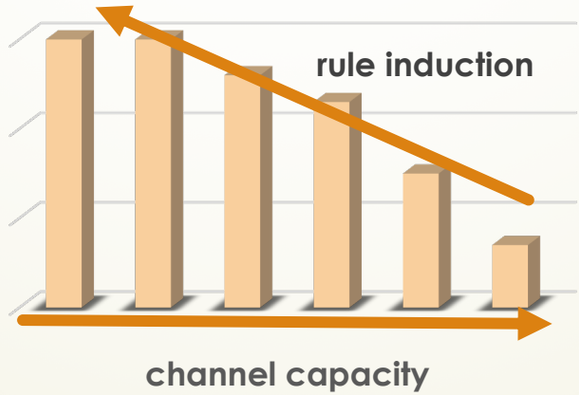
Entropy Model - hypotheses

Rule Induction → interaction of ***input complexity*** **(entropy)** and ***channel capacity***

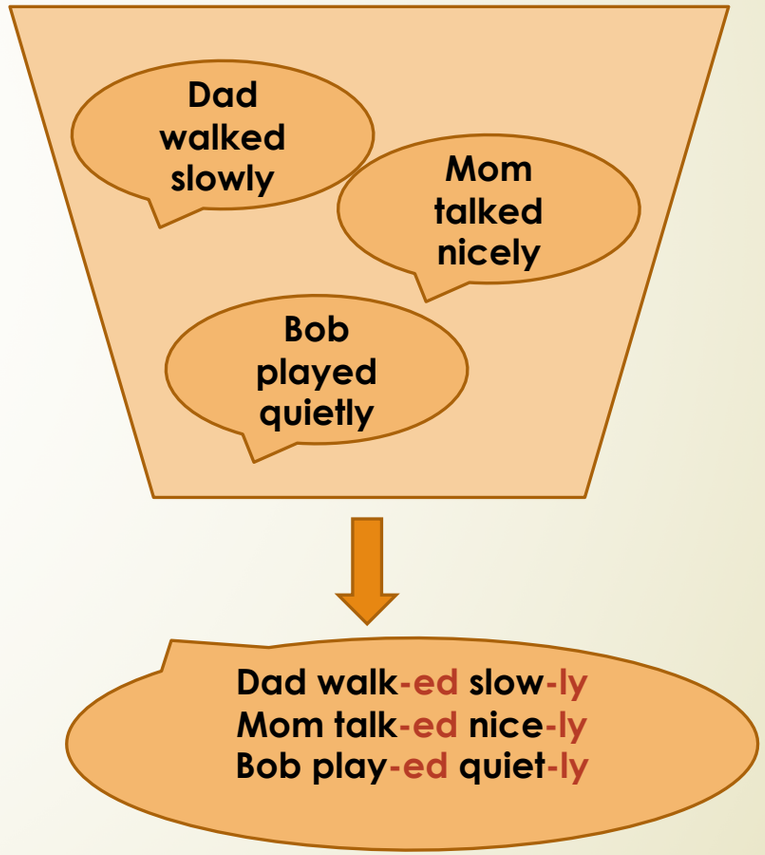
Low channel capacity →
category-based generalizations







High channel capacity →
item-bound generalizations



Effect of Channel Capacity on Rule Induction

Channel Capacity

- information-theoretic concept (entropy/time)
- model the limited encoding power of the brain

? What are the cognitive processes that modulate channel capacity? ?

- memory capacity
- attention
- pattern-recognition capacity
- …

? How to tamper with channel capacity to boost rule induction ?

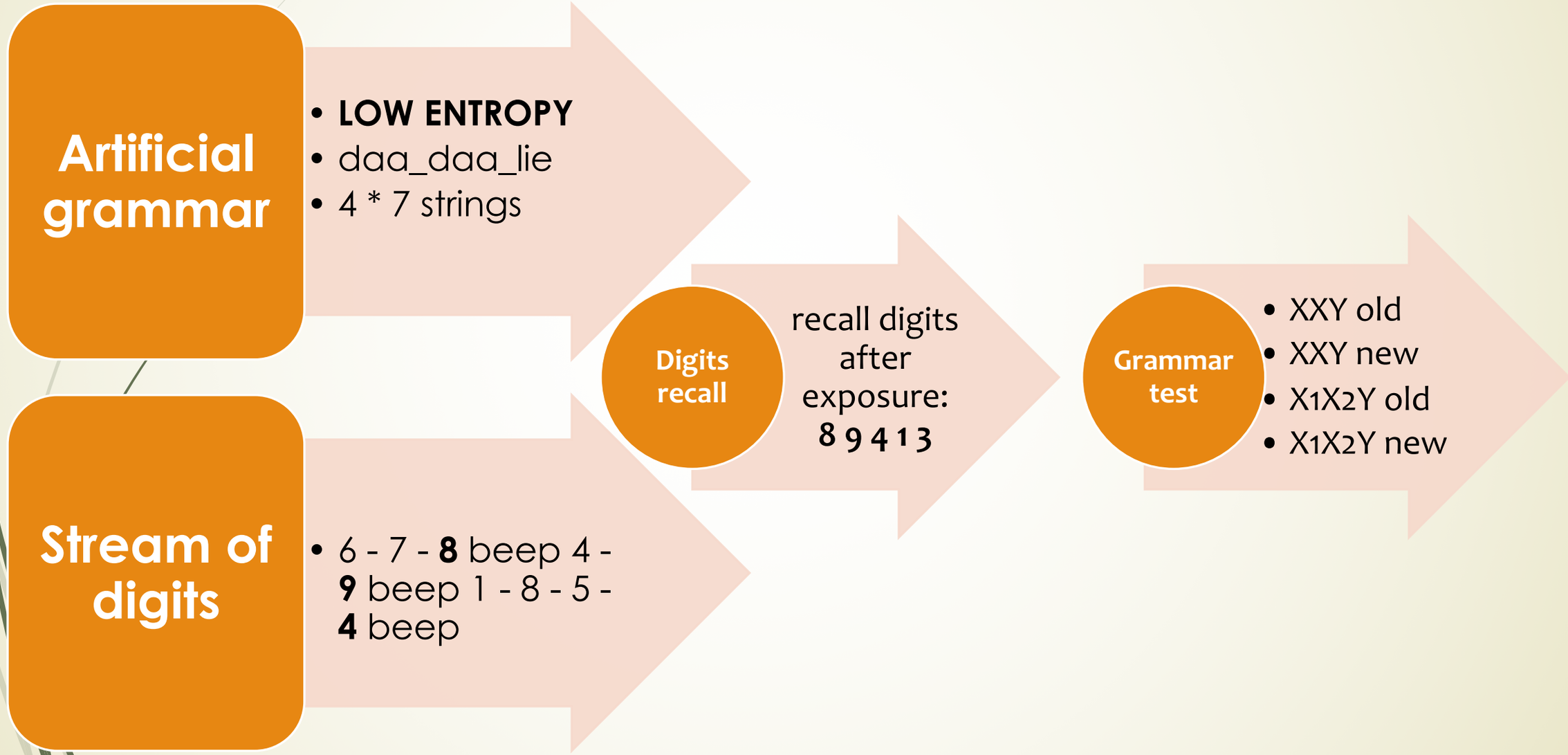
→ keep **Input complexity** low: 4*7 X/4*7 Y (2.8 bits)
→ exceed **Channel Capacity**

- **Experiment 3: 88 adults (age 19 -42)**
  - **same AGL - 3-syllable XXY: *daa_daa_lie***
- **tamper with memory capacity & attention**
  - **working memory (similar to a running memory test)**
  - **attention (irrelevant stream of digits in background)**
  - **pre-loaded memory (adjusted to individual span)**

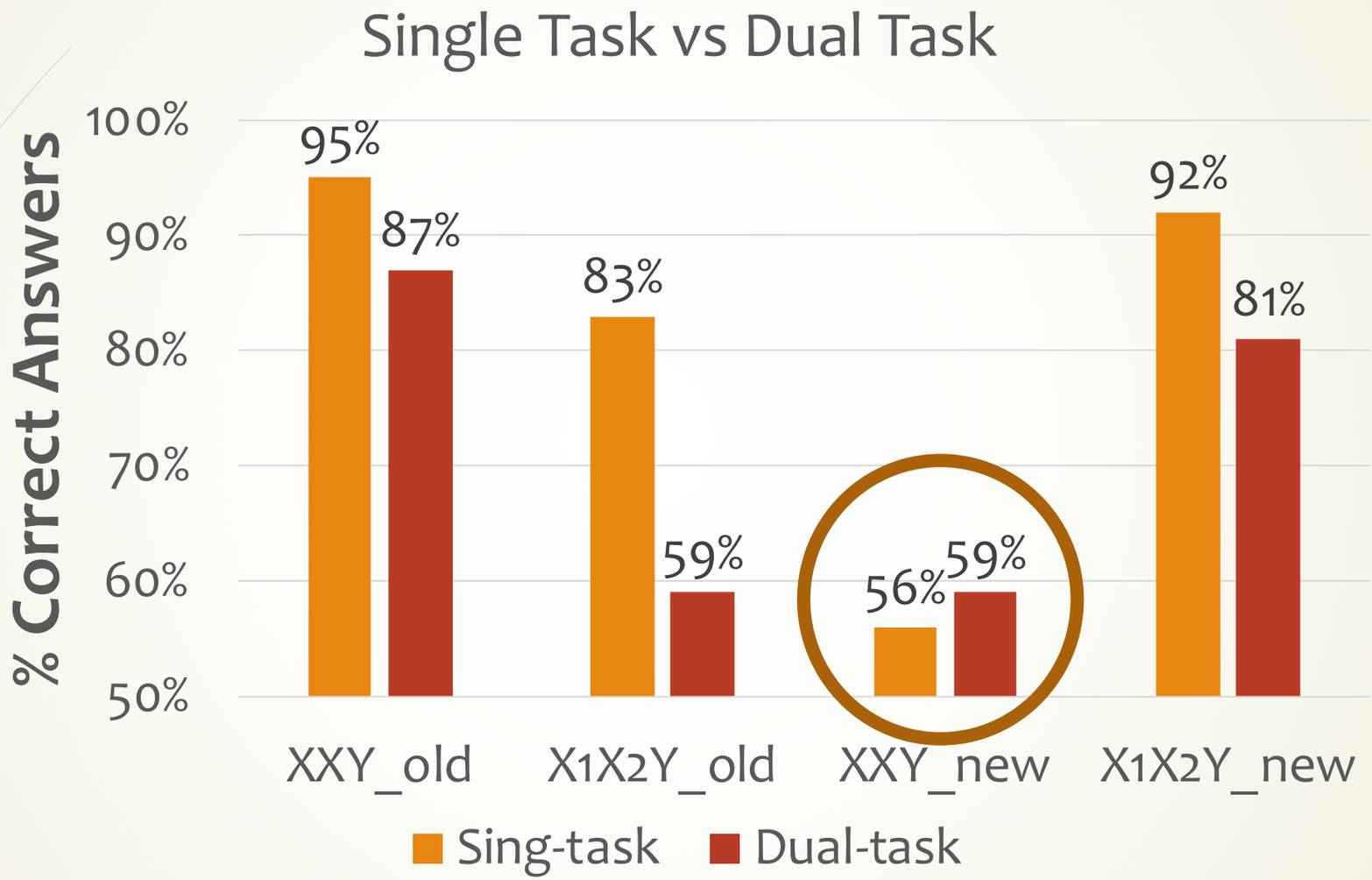
Experimental Conditions

1. Dual-task
2. Distractor
3. Pre-load

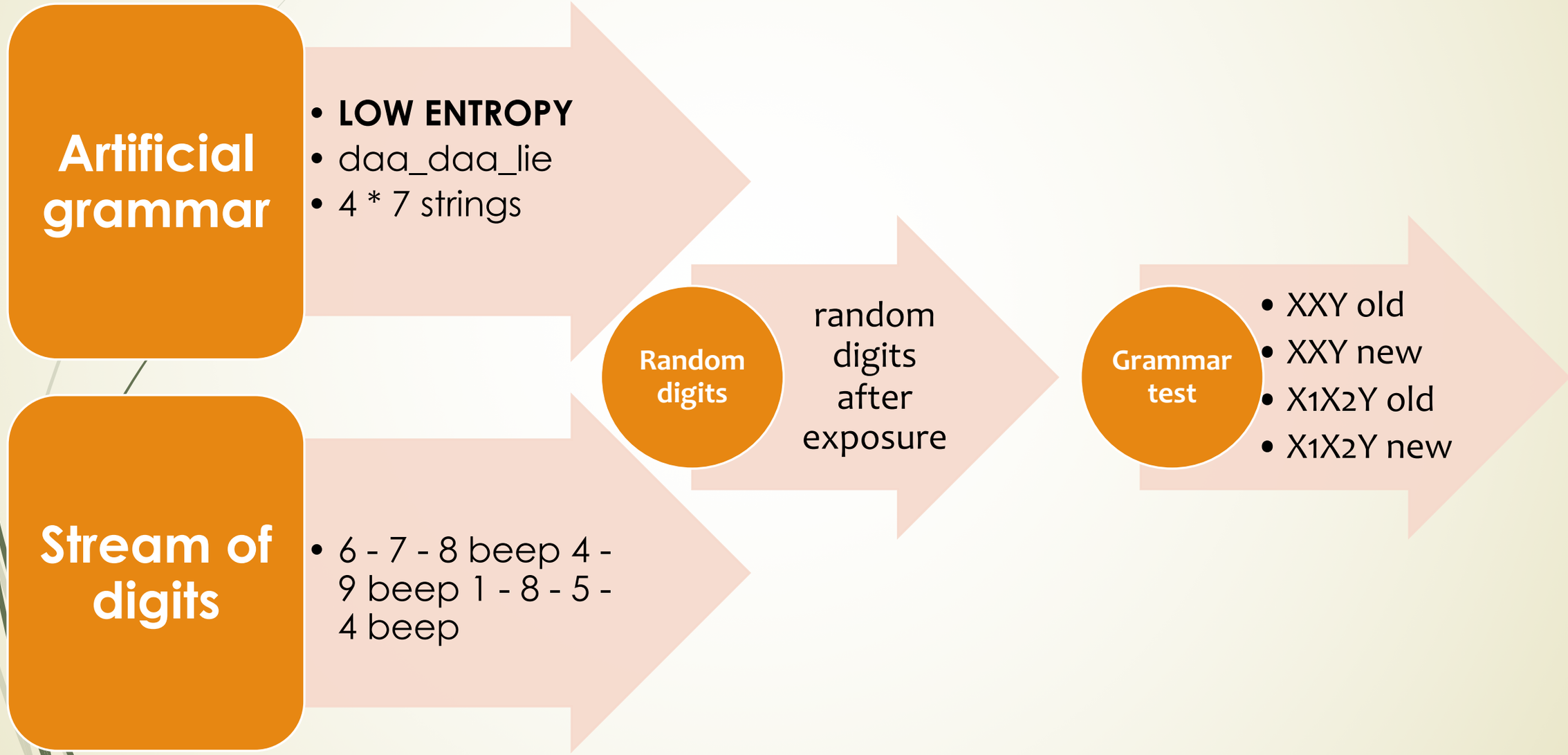
Dual-task: timeline



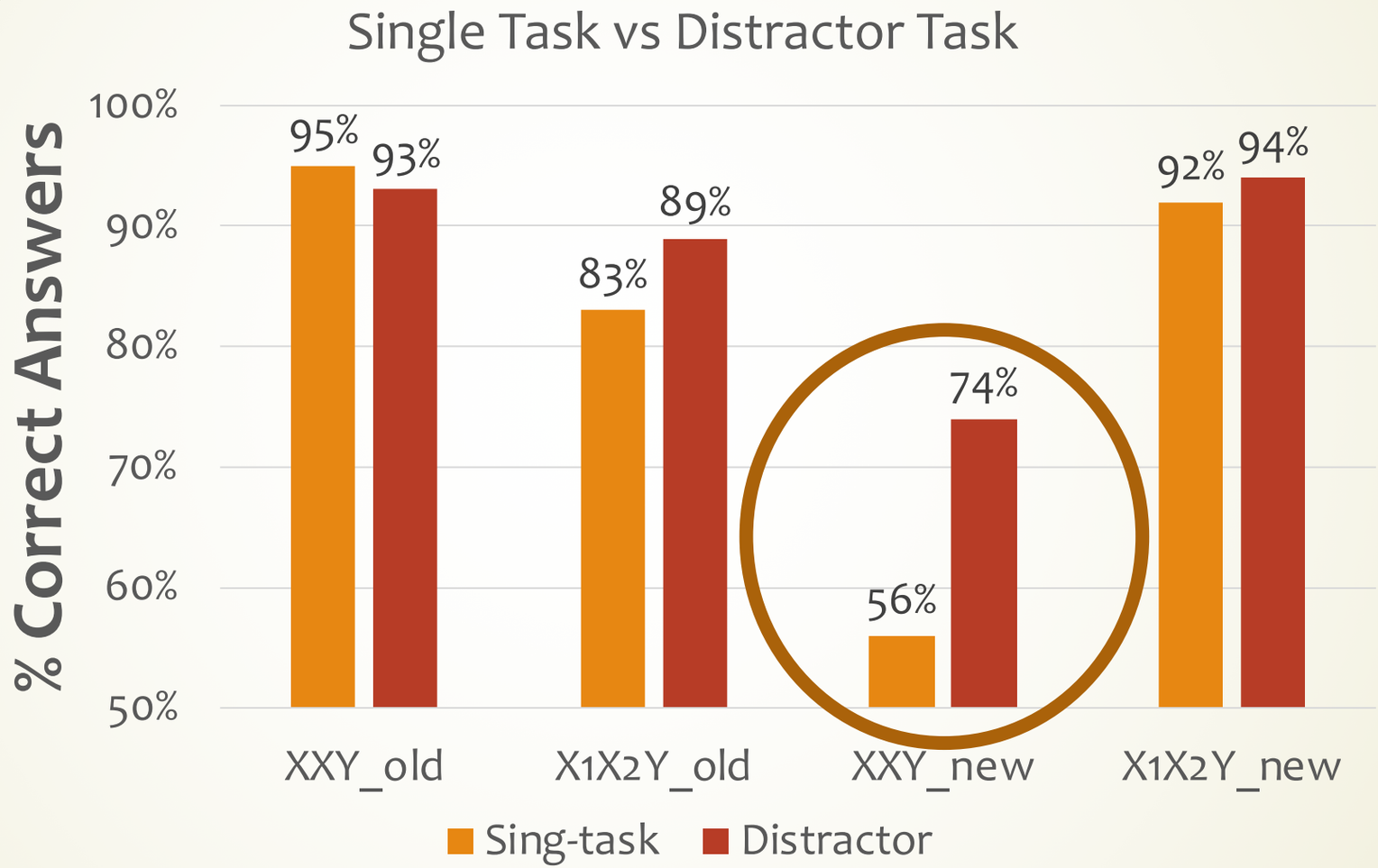
Results: Dual-task



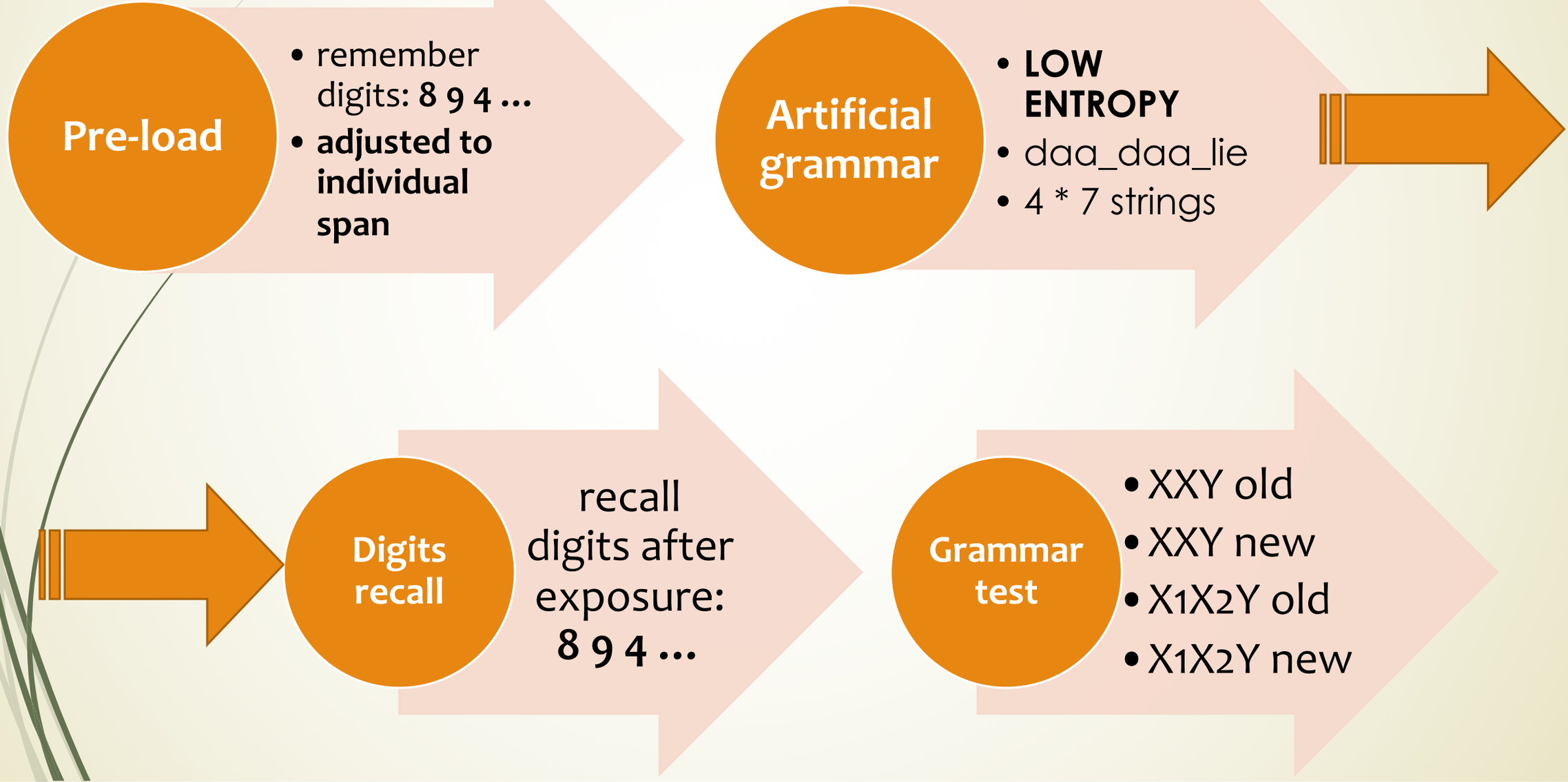
Distractor: timeline



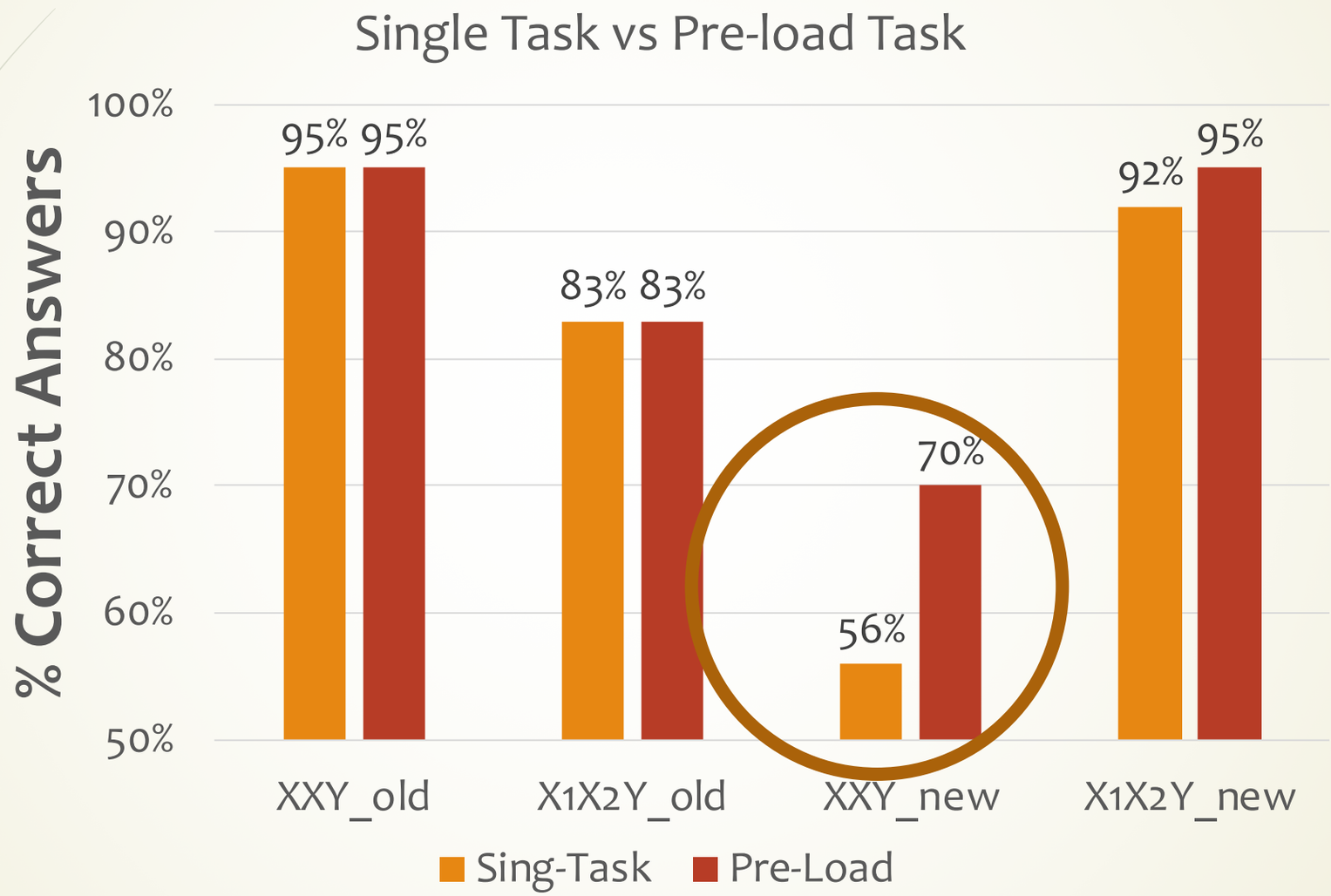
Results: Distractor



Pre-load: timeline



Results: Pre-load

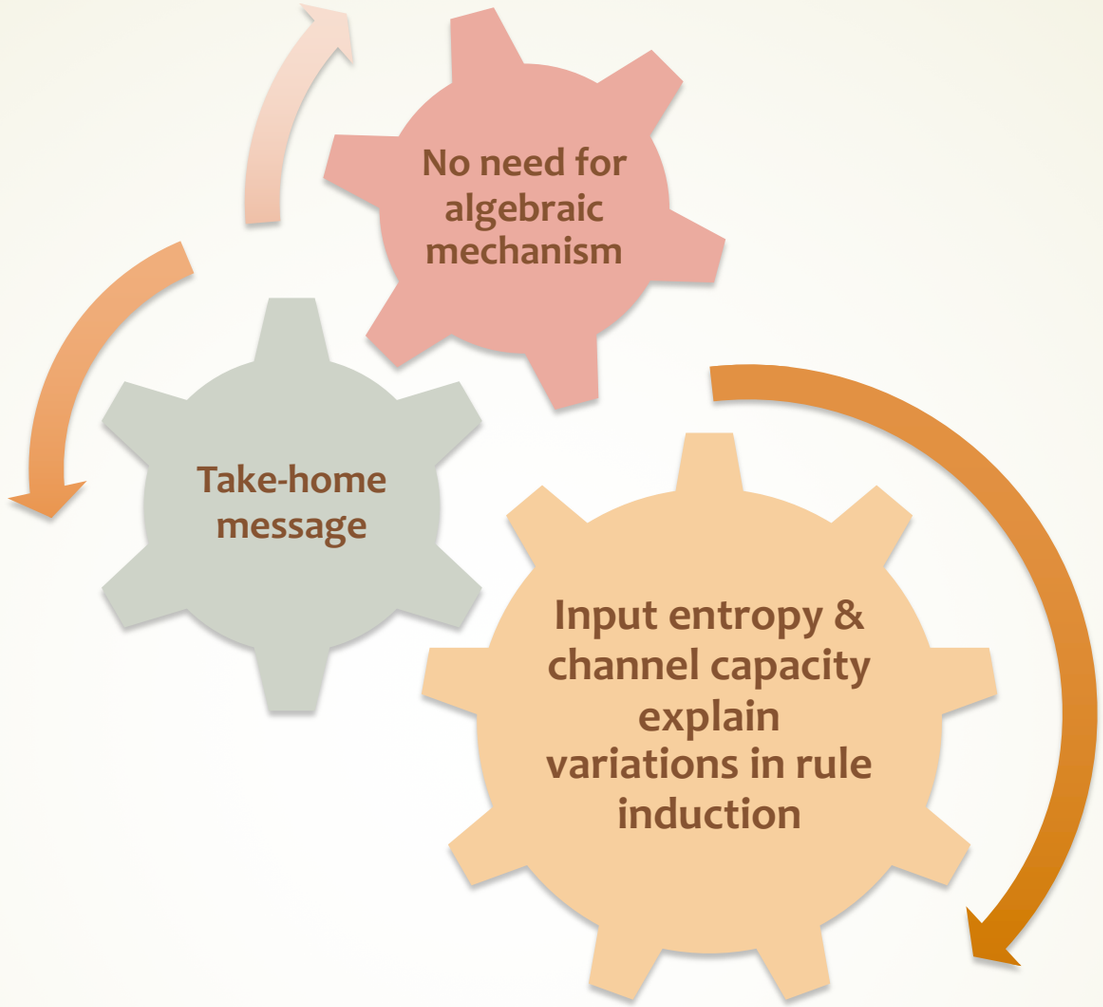


Conclusions

→ if input entropy increases, the tendency to generalize increases gradually

→ exceeding channel capacity leads to a higher tendency to generalize

→ the two types of rule induction are **outcomes of the same information encoding mechanism**
→ *gradually* moves from lower-level *item-bound* encoding to higher-level *category-based* encoding in response to the interaction between **input entropy** and the encoding power (**channel capacity**)



Utrecht University

Netherlands Organisation for Scientific Research

UiL OTS
UTRECHT INSTITUTE OF LINGUISTICS

"To think is to forget a difference, to generalize, to abstract. In the overly replete world of Funes there were nothing but details, almost contiguous details."

Funes, The Memorious

Jorge Luis Borges