Universiteit Utrecht

NWO Netherlands Organisation for Scientific Research

UiL OTS Utrecht Institute of Linguistics

Statistical Learning and Cognitive Constraints on Rule Induction
An Entropy Model

Silvia RĂDULESCU, Efi GIANNOPOULOU, Frank WIJNEN, Sergey AVRUTIN

From little evidence to abstract rules in language acquisition

(1) statistical learning (Aslin & Newport, 2012)
(2) algebra-like system (Marcus et al, 1999)

An Entropy Model

Entropy (input complexity) → Channel capacity (encoding power = entropy/time)

Entropy → a function of the number of different items in the input and their probability of occurrence (frequency)
→ a measure of input complexity (bits)

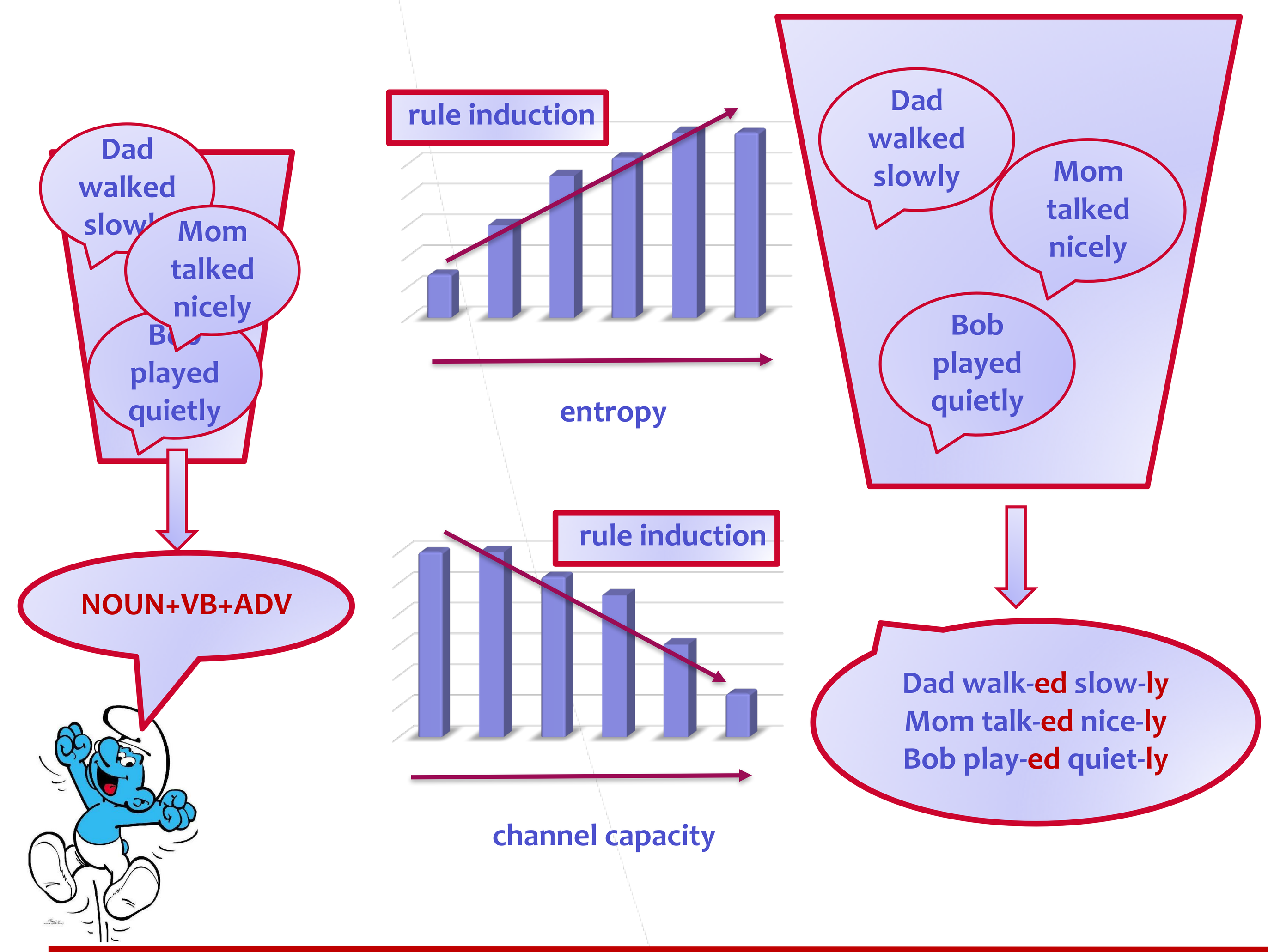
$$H(X) = -\sum_{i=1}^{n} p(x_i) log p(x_i)$$

(Shannon, 1948)

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

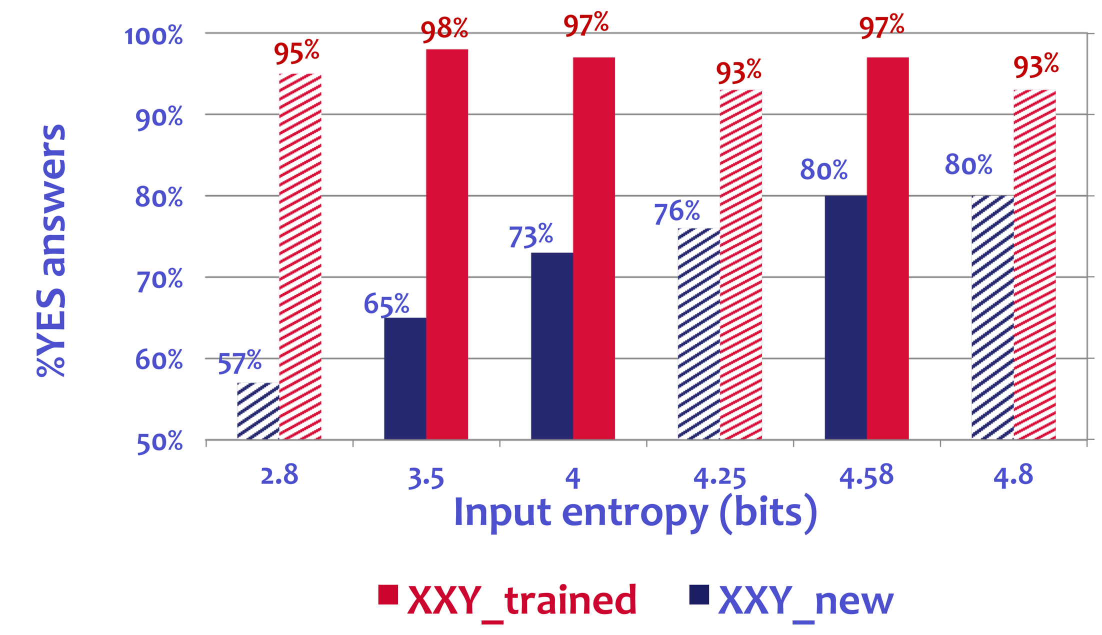
Entropy > channel capacity → category-based generalizations

Entropy < channel capacity → item-bound generalizations



Experiments 1&2 - Effect of Entropy on Rule Induction

- 71 adults, ~22y, ~4min, between-subjects
- 3-syllable XXY: *goo_goo_sjie*
- manipulated *ENTROPY* (number & frequency)
  - 2.8 bits (4 × 7Xs / 4 × 7Ys)
  - 3.5 bits (4 × 6Xs / 4 × 6Ys)
  - 4 bits (2 × 12Xs / 2 × 12Ys)
  - 4.25 bits (2 × 14Xs / 2 × 14Ys)
  - 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?")

5x4=20 items

- ➢ XXY_trained_syllables: goo_goo_sjie √
- ➢ X_1X_2Y_new_syllables: reu_loo_gee *
- ➢ XXY_new_syllables: too_too_suu √
- ➢ X_1X_2Y _trained_syllables: teu_duu_saa *

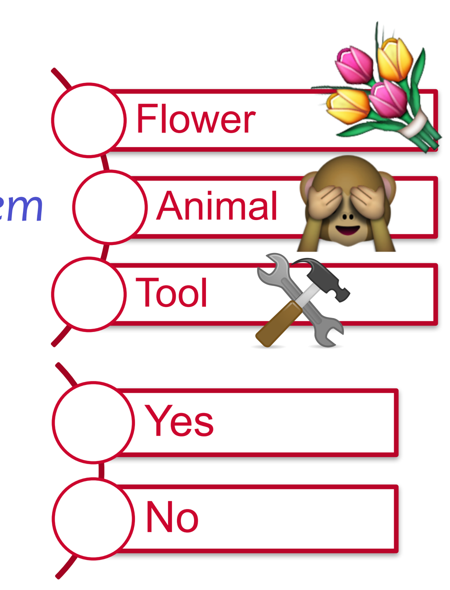
Yes / No

Experiment 3 - Effect of Channel Capacity on Rule Induction

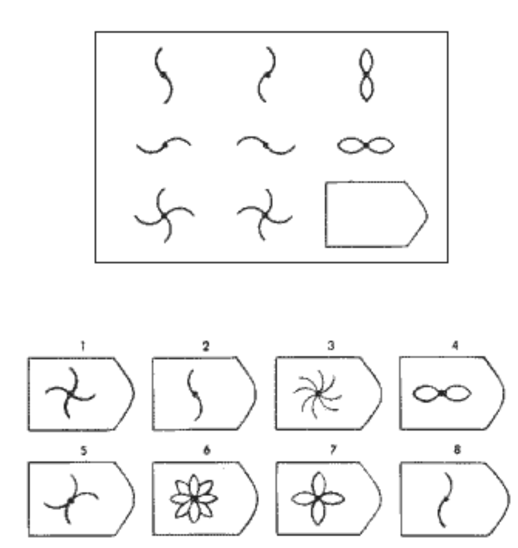
- ➢ 51 adults (age 19-44)
- ➢ Medium Entropy: 2*14 X/2*14 Y (4.2 bits)
- ➢ 3 independent tasks: Forward Digit Span (**FDS**), Incidental Memorization Task (**IMT**), Raven's Standard Progressive Matrices (**RSPM**)

Incidental Memorization Task

- **Training:**
  - ➢ 30 non-sense bi-syllabic words: *go_pem*
  - ➢ What does this word sound like?
- **Surprise memory test:**
  - ➢ Have you heard this word before?
    - ➢ 13 targets + 13 foils

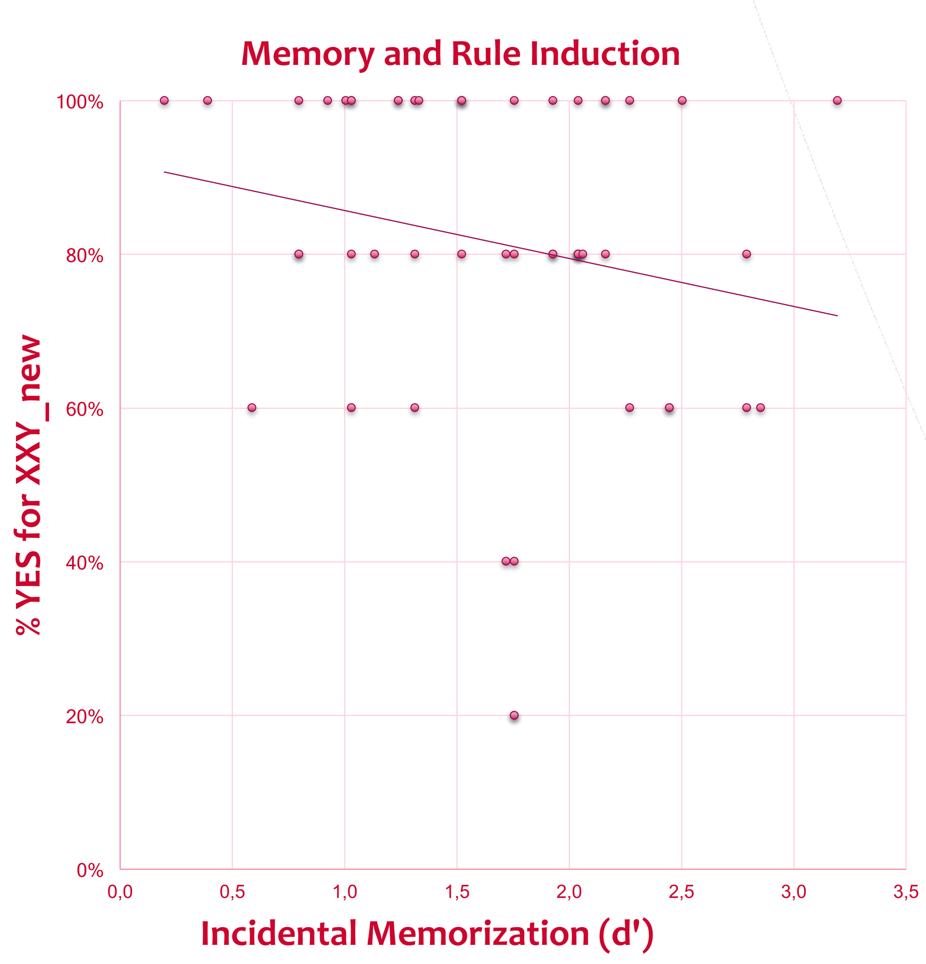


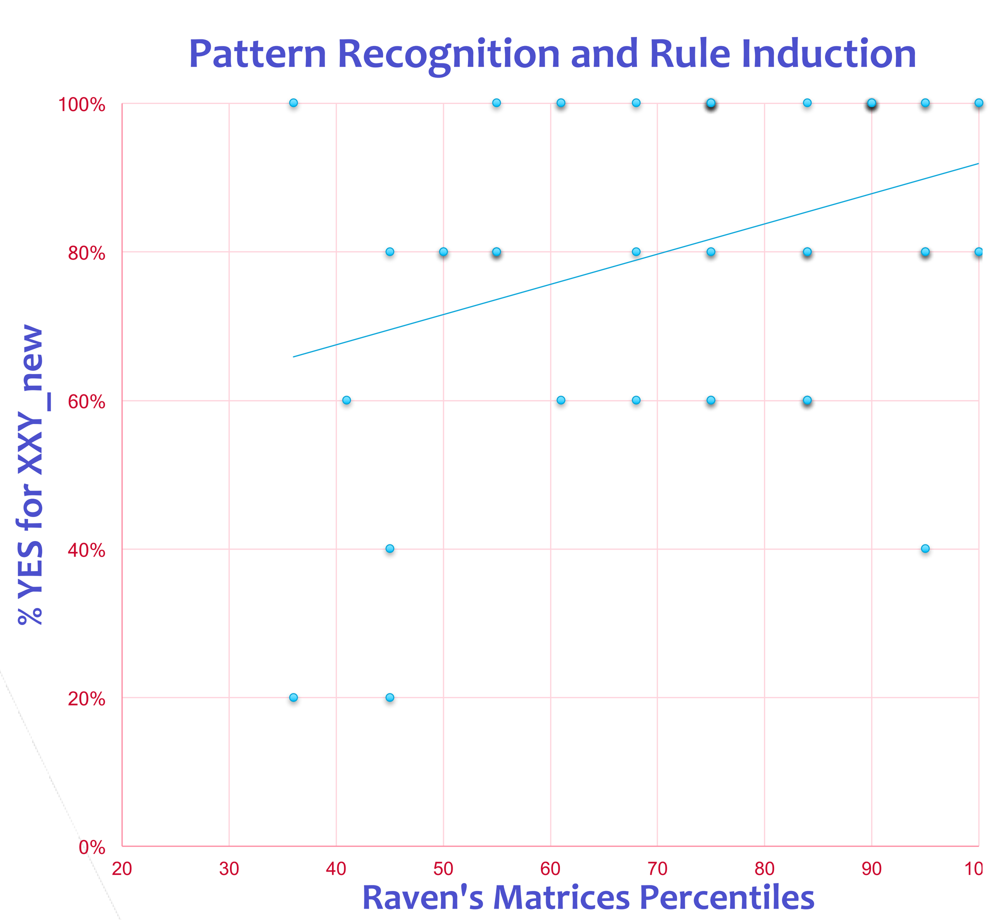
Raven's Standard Progressive Matrices

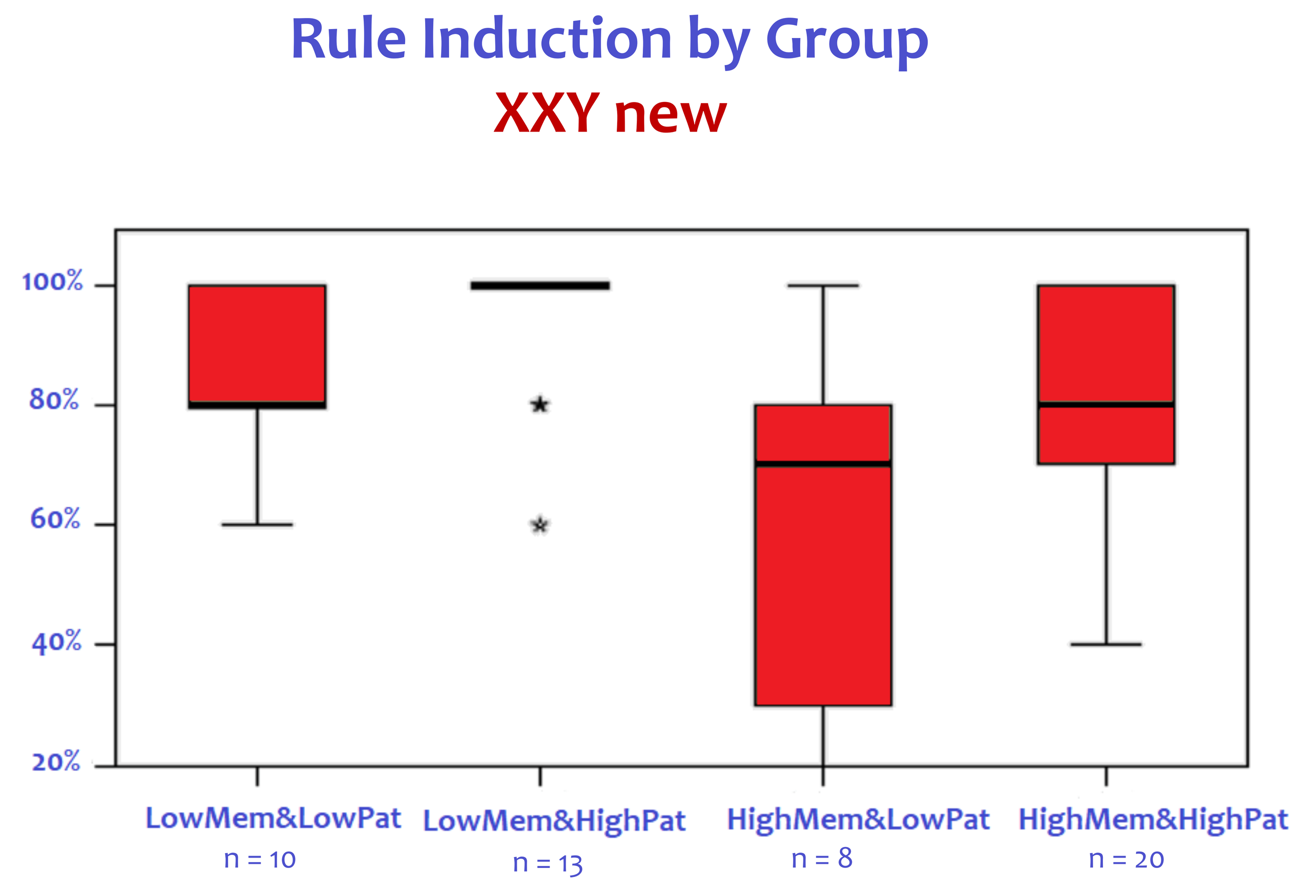


Results

Ordinal Regression. Covariates: scores on the tests (FDS, IMT, RSPM). RSPM: significant **positive** effect on XXY_new and X_1X_2Y _trained; IMT: significant **negative** effect on XXY_new and X_1X_2Y _trained.







Discussion

Rule induction -> the interaction between input entropy and a limited encoding power of the brain. A low entropy in the input does not boost generalization *per se*, so it allows for more variation in participants' individual tendencies to generalize. Thus incidental memory and pattern recognition are predicted to better explain that variation.

No need for algebraic rules. Cognitive constraints on statistical learning explain variations in rule induction.

Conclusions

If input entropy increases, the tendency to generalize increases gradually.
Lower incidental memory predicts a higher tendency to generalize.
Higher visual pattern recognition predicts a higher tendency to generalize.

References

Aslin, R.N., and Newport, E. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science*, 21, 170–176.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80.

Shannon, C. E. (1948). *A mathematical theory of communication*. Bell System Technical Journal, 27, 379–423.