Bigram coding as a general visual mechanism

(Nothing special about reading?)

Davide Crepaldi

International School for Advanced Studies (SISSA), Trieste

https://lrlac.sissa.it
Tweets at @CrepaldiDavide

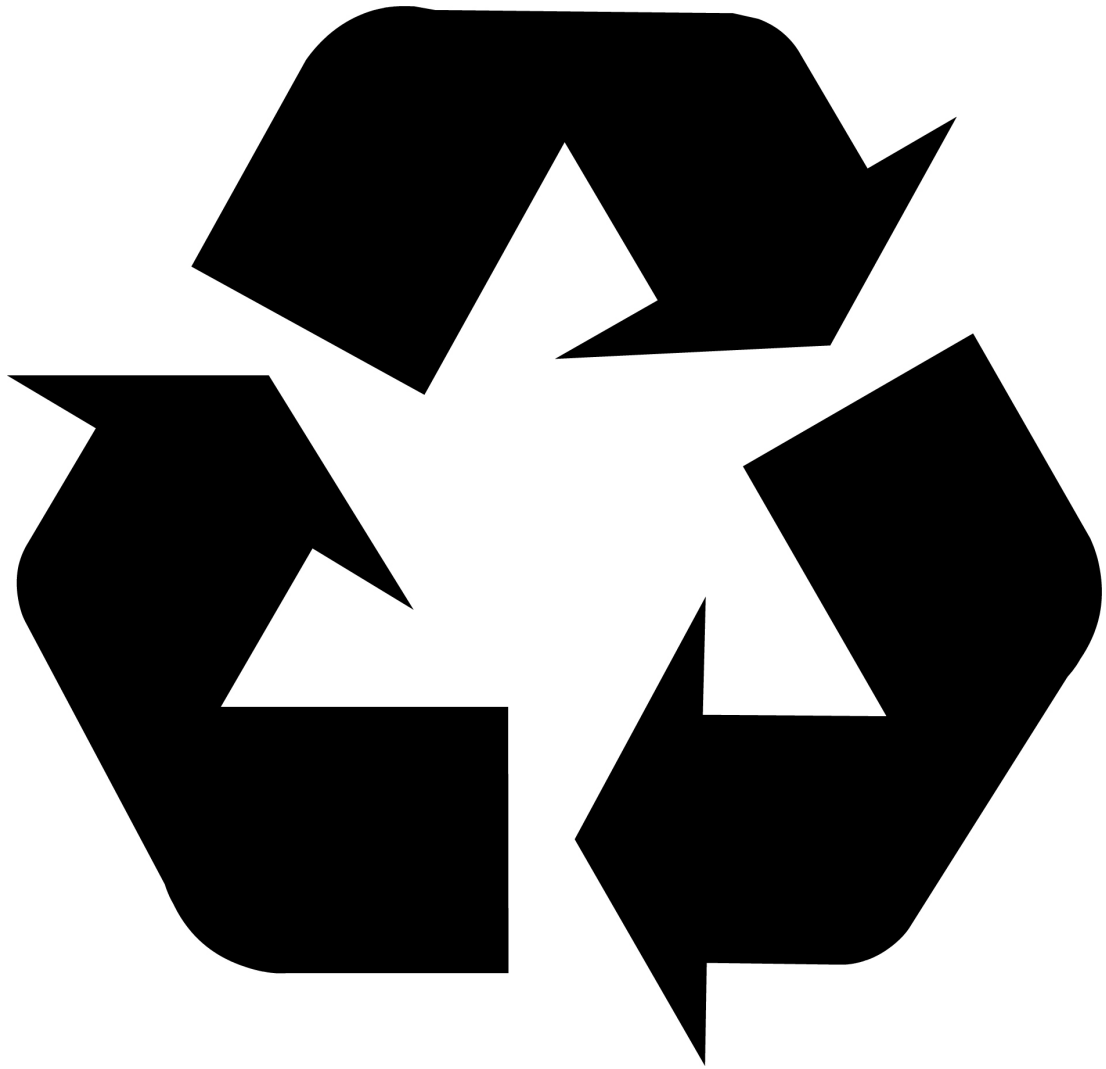
ESCoP, Tenerife, 28 September 2019

Reading is a human wonder

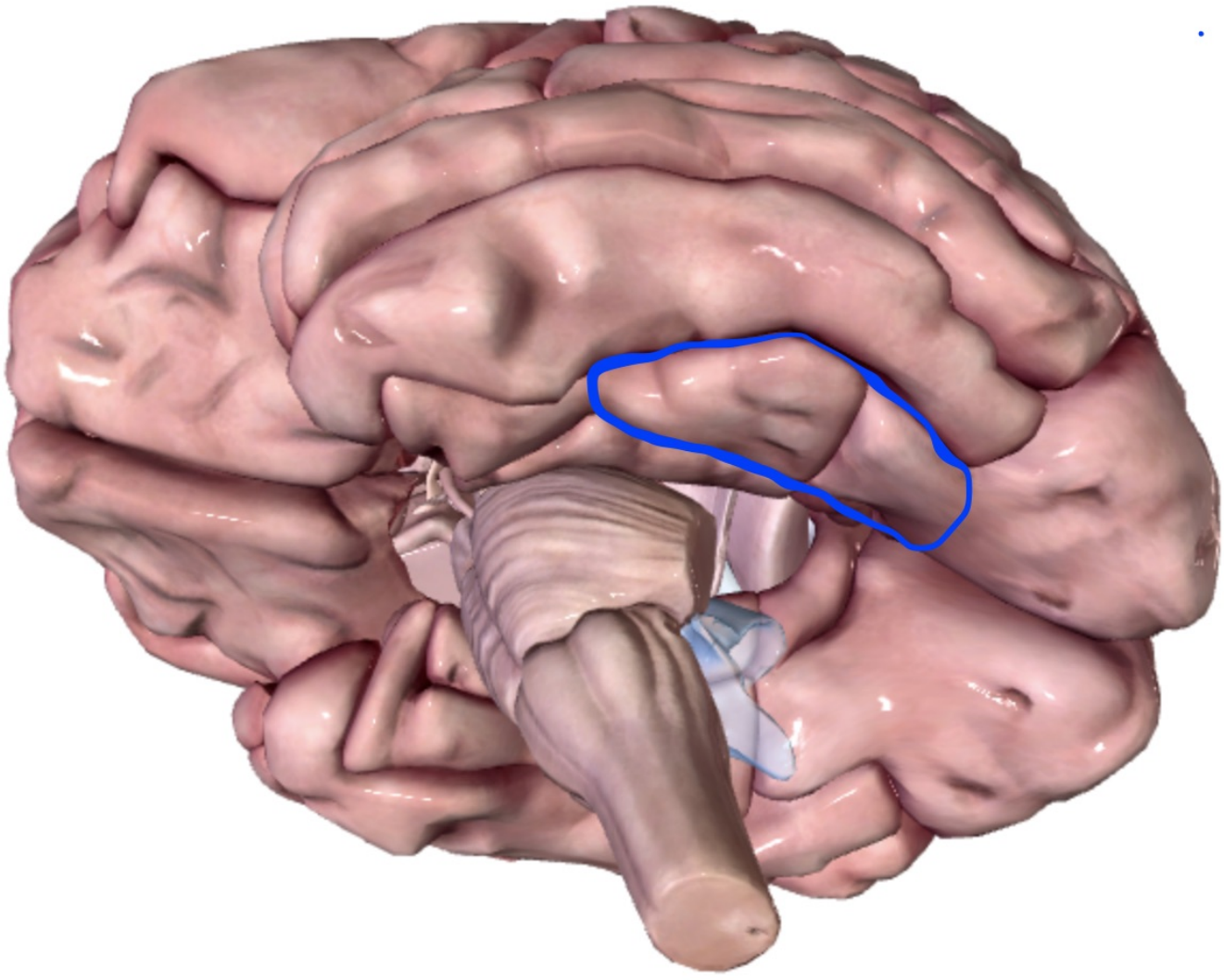
- We can read 8–letter words in ~35ms (e.g., Forster and Davis, 1984)
- We gather information about ~20 letters every ~200ms (e.g., Rayner, 1998)
- We read ~250 words per minute (e.g., Brysbaert, 2019)

No genetic endowment

- Written language isn't observed universally
- Literacy isn't acquired spontaneously
- Writing appeared ~5.5K years ago

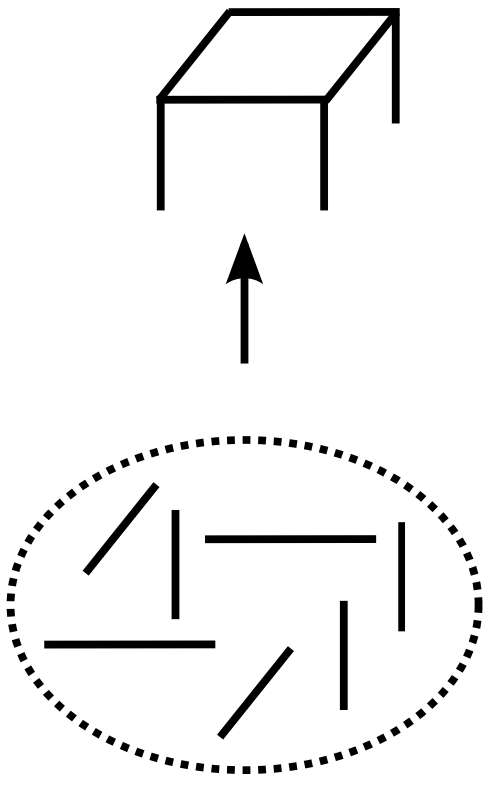


The ventral occipito–temporal cortex (VOTC)



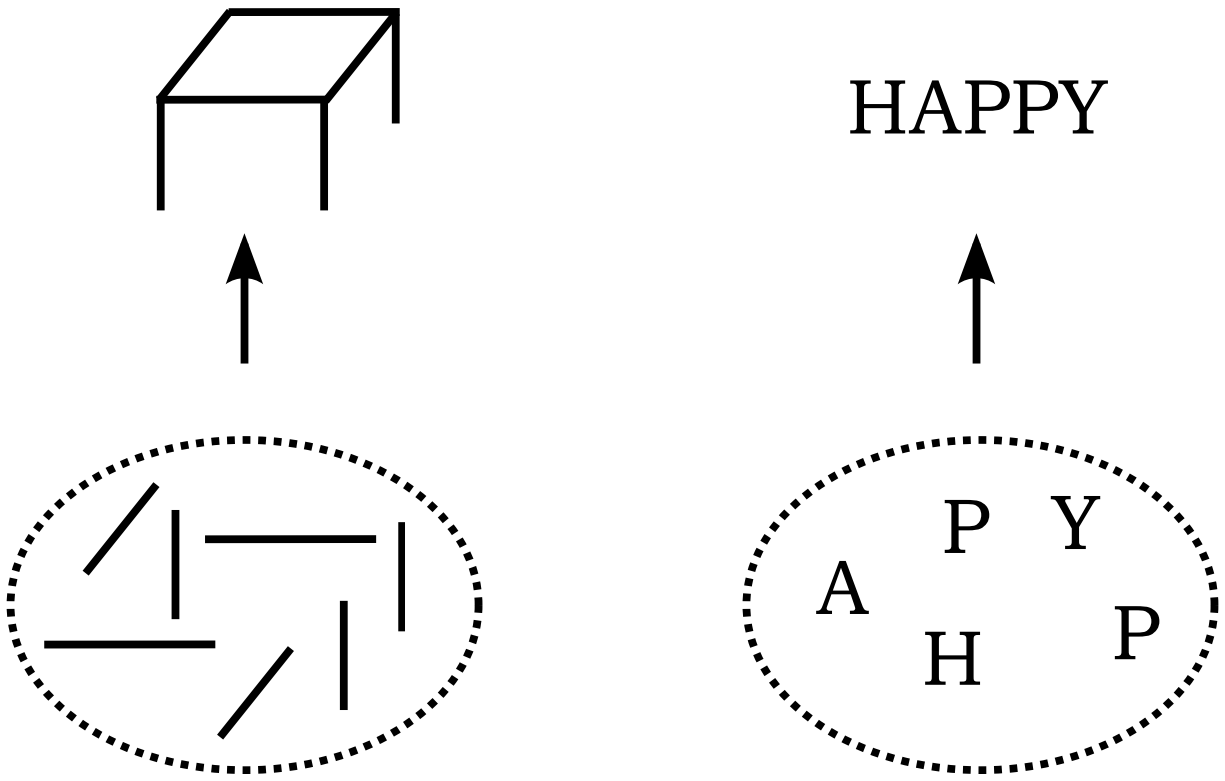
The Statistical Learning hypothesis

- Find regularities in low–level units
- Build higher–level units based on these regularities



The Statistical Learning hypothesis

- Find regularities in low–level units
- Build higher–level units based on these regularities



Statistical Learning and reading?

Nah

- Reading ability doesn't really correlate with statistical learning tasks (Schmalz et al., 2019)
- No bigram frequency effects in lexical decision (Schmalz and Mulatti, 2017)

Yeah!

- Orthographic processing in non–linguistic animals (Grainger et al., 2012; Rajalingham et al., 2019)
- Sensitivity to frequent bigrams in artificial scripts (Chetail, 2017)

Phantom words paradigm (Endress and Mehler, 2009)

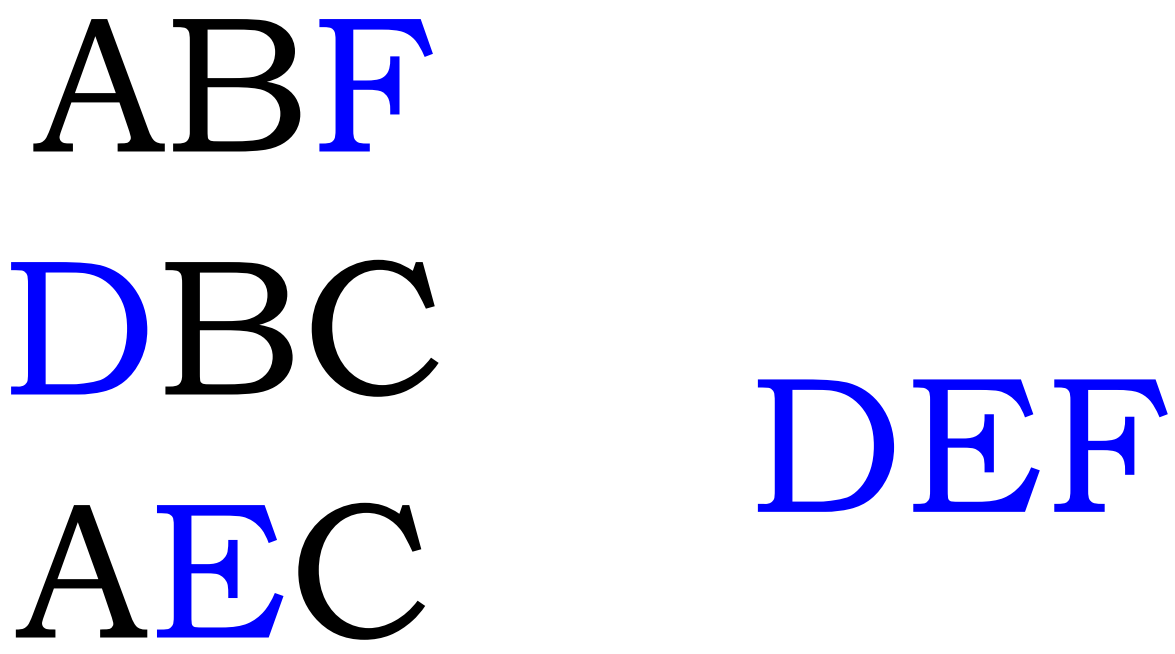
ABF

DBC

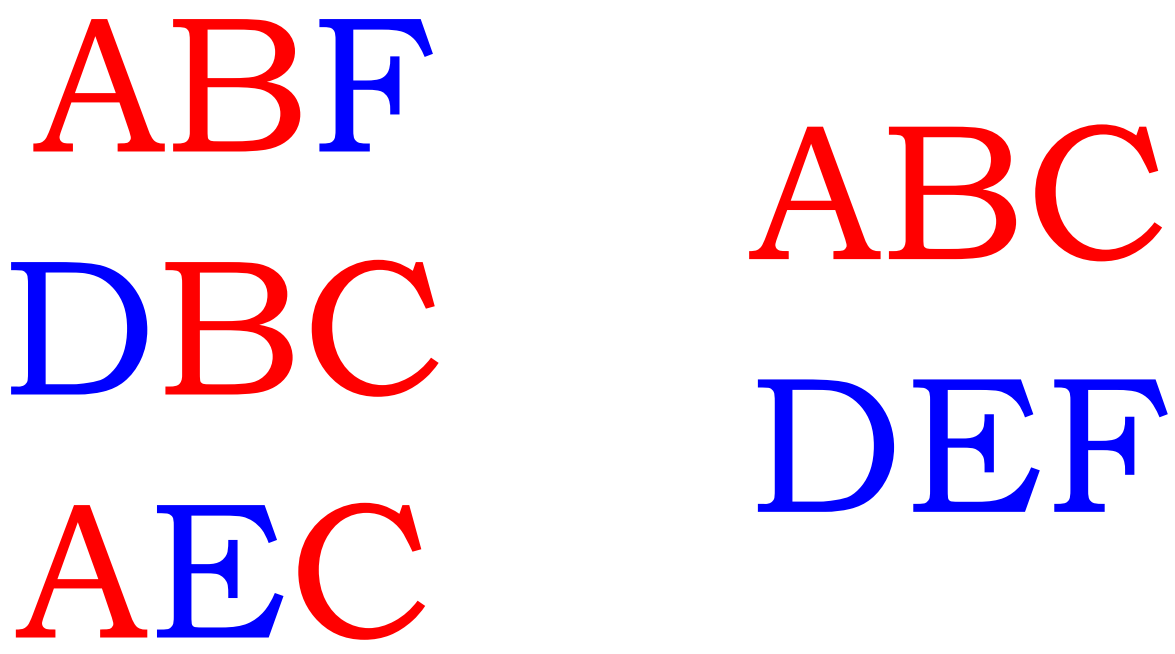
AEC

Phantom words paradigm (Endress and Mehler, 2009)

Phantom words paradigm (Endress and Mehler, 2009)

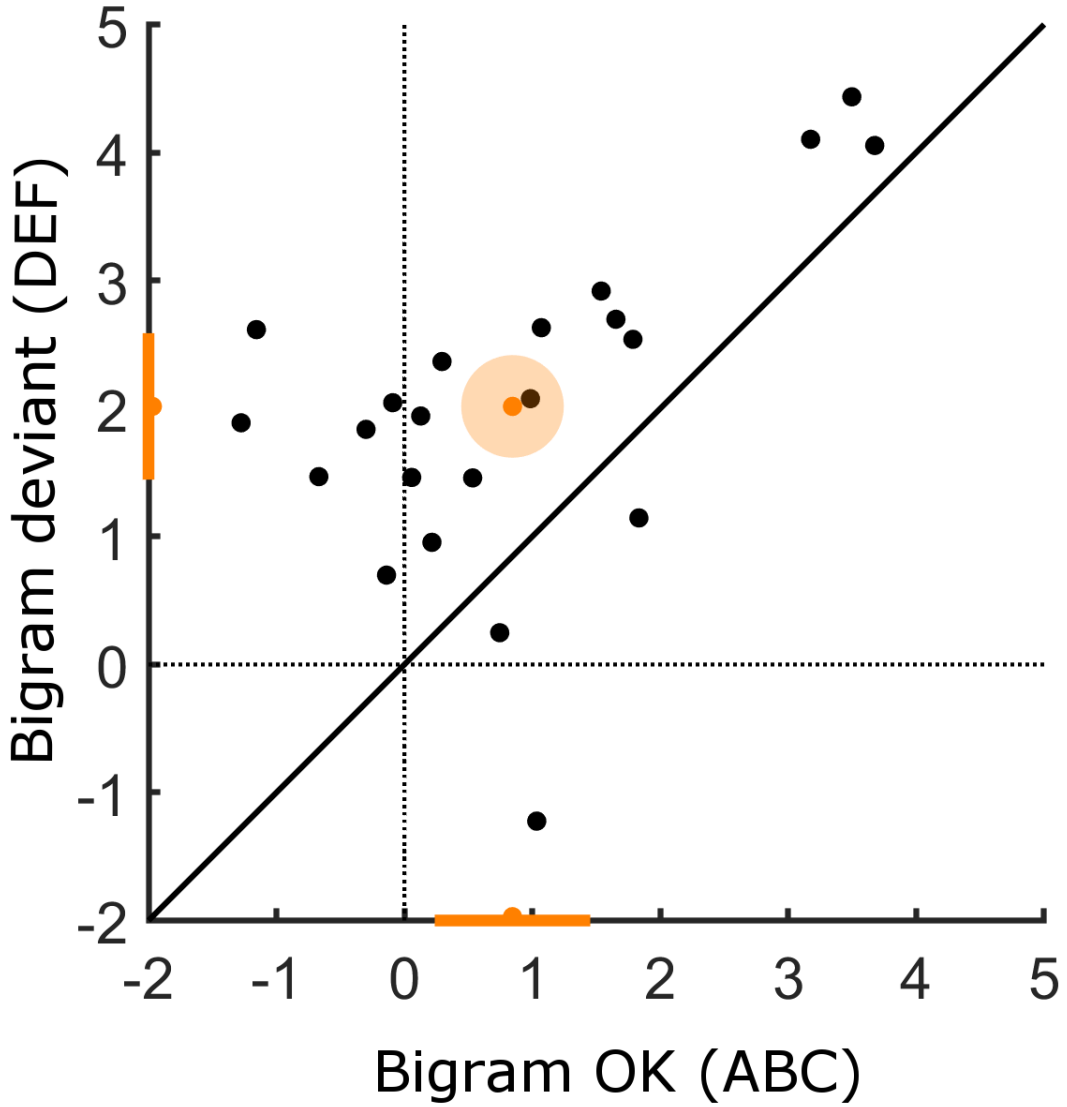


Phantom words paradigm (Endress and Mehler, 2009)

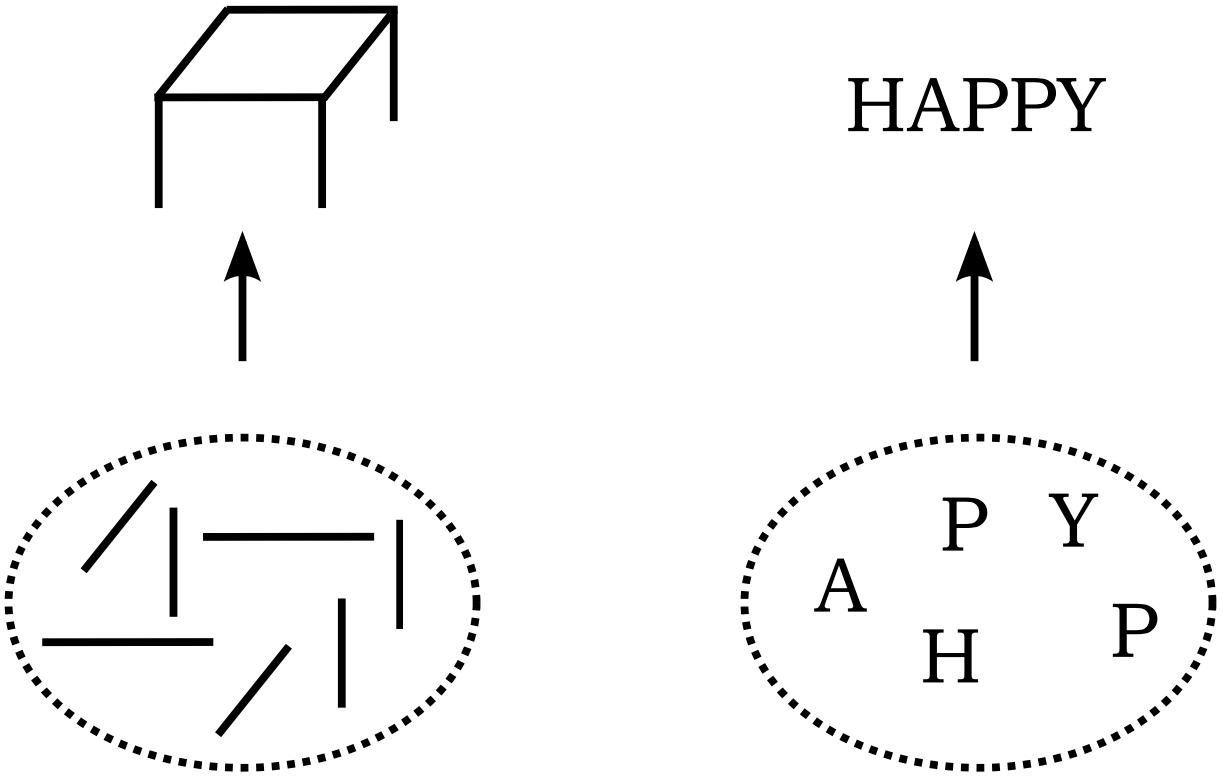


Pseudo–letters (Vidal et al., 2017)

Results

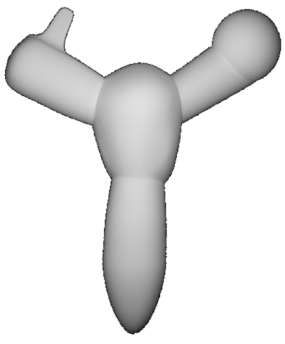
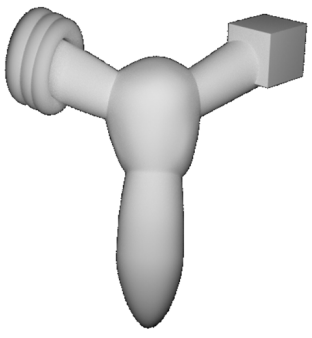
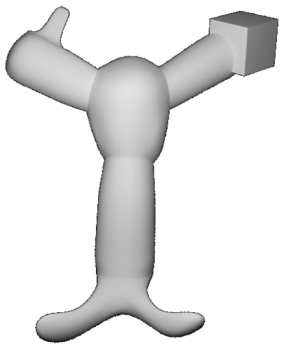
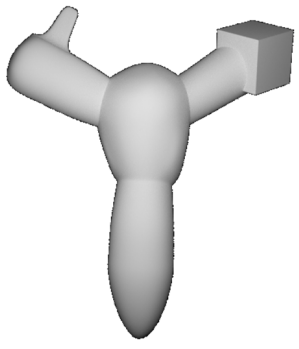
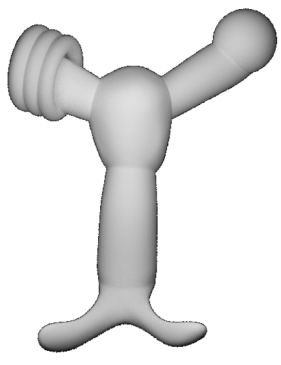


The Statistical Learning hypothesis

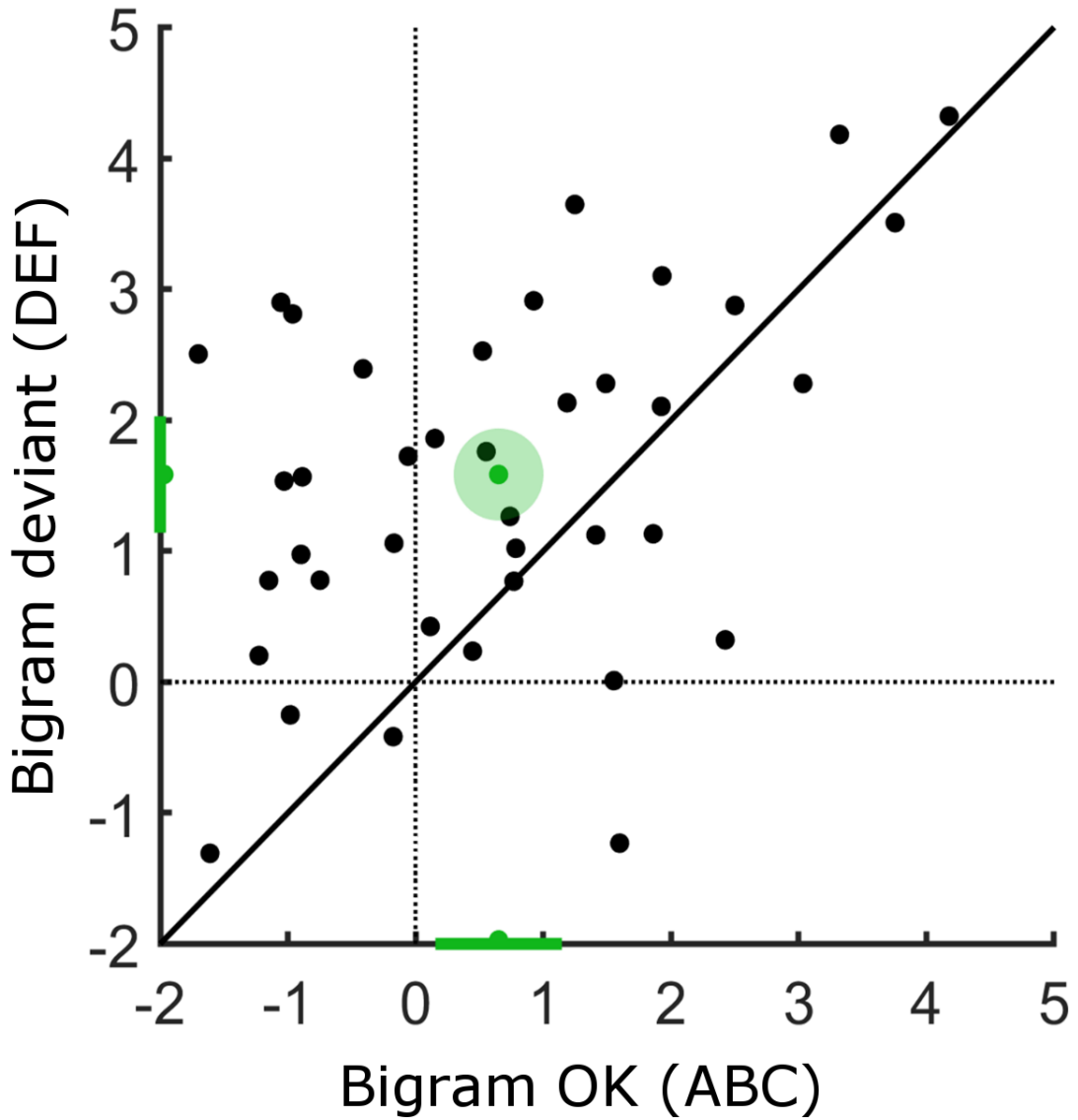


Vision, not language

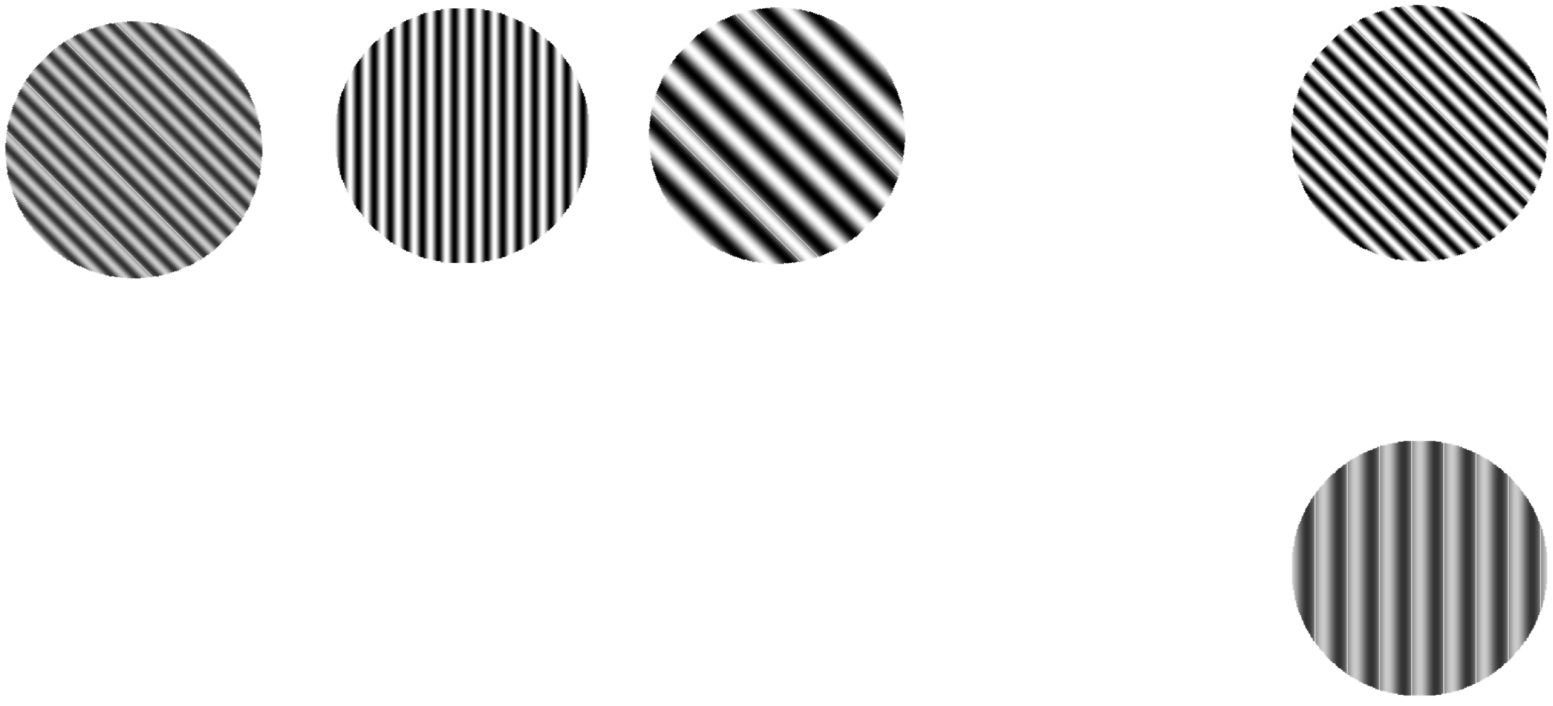
Phantom tripods

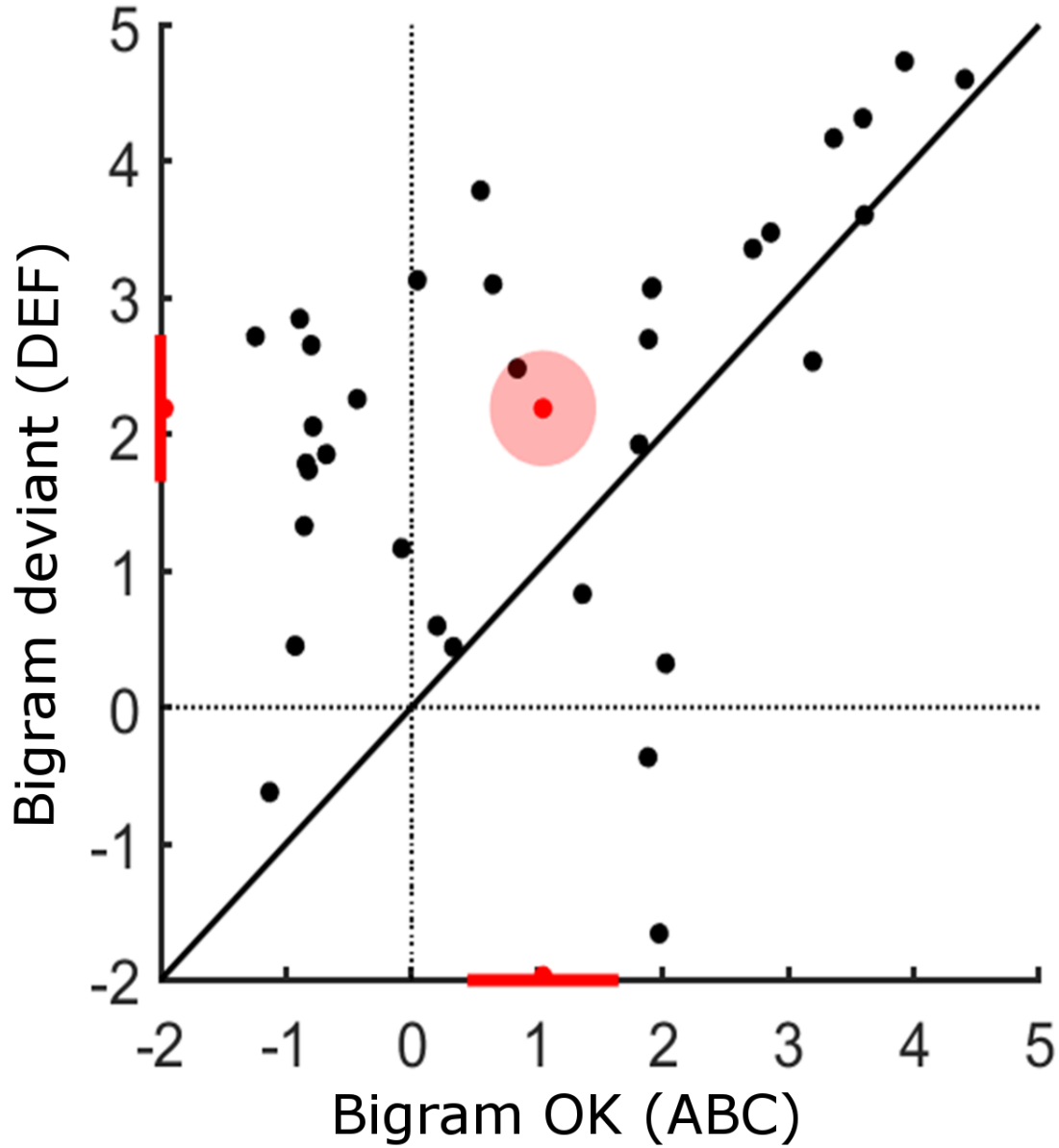
Phantom tripods



Phantom Gabors



Phantom Gabors



Take–home message

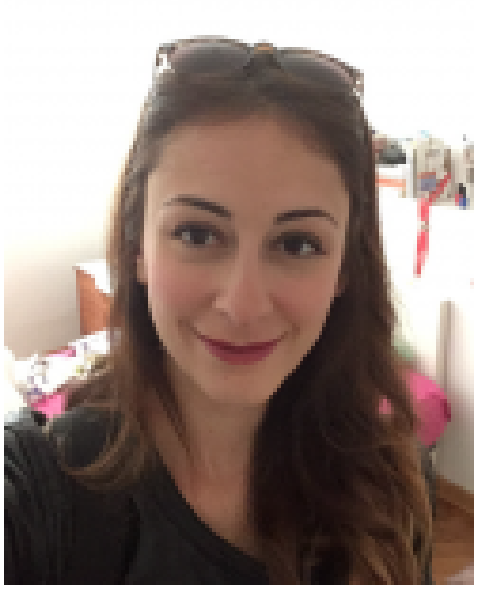
- We code for nGrams/letter transition stats while learning novel words
- We use the same mechanism while learning novel objects, where the lower–level units are:
  - not arranged horizontally, and very different visually from letters
  - not even spatially segregated

- More generally, reading shares (part of) its computational core with vision
- This computational core is captured by the statistical learning hypothesis

Acknowledgments

Yamil Vidal

Eva Viviani



StG 2015 – STATLEARN

Bigram coding as a general visual mechanism

(Nothing special about reading?)

Davide Crepaldi

International School for Advanced Studies (SISSA), Trieste

https://lrlac.sissa.it
Tweets at @CrepaldiDavide

ESCoP, Tenerife, 28 September 2019

References I

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, 109.

Chetail, F. (2017). What do we do with what we learn? statistical learning of orthographic regularities impacts written word processing. *Cognition*, 163:103–120.

Endress, A. D. and Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language*, 60:351–367.

Forster, K. I. and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning Memory and Cognition*, 10:680–698.

Grainger, J., Dufau, S., Montant, M., Ziegler, J., and Fagot, J. (2012). Orthographic processing in baboons (papio papio). *Science*, 336(6078):245–248.

Rajalingham, R., Kar, K., Sanghavi, S., Dehaene, S., and DiCarlo, J. J. (2019). A potential cortical precursor of visual word form recognition in untrained monkeys. *bioRxiv*.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124:372–422.

References II

Schmalz, X., Moll, K., Mulatti, C., and Schulte-Körne, G. (2019). Is statistical learning related to reading ability, and if so, why? *Scientific Studies of Reading*, 23:64–76.

Schmalz, X. and Mulatti, C. (2017). Busting a myth with the bayes factor: Effects of letter bigram frequency in visual lexical decision do not reflect reading processes. *The Mental Lexicon*, 12:263–282. Retrieved from osf.io/3ybwd.

Vidal, C., Content, A., and Chetail, F. (2017). BACS: The Brussel Artificial Character Set for studies in Cognitive Psychology and Neuroscience. *Behavior Research Methods*, 49:2093–2112.

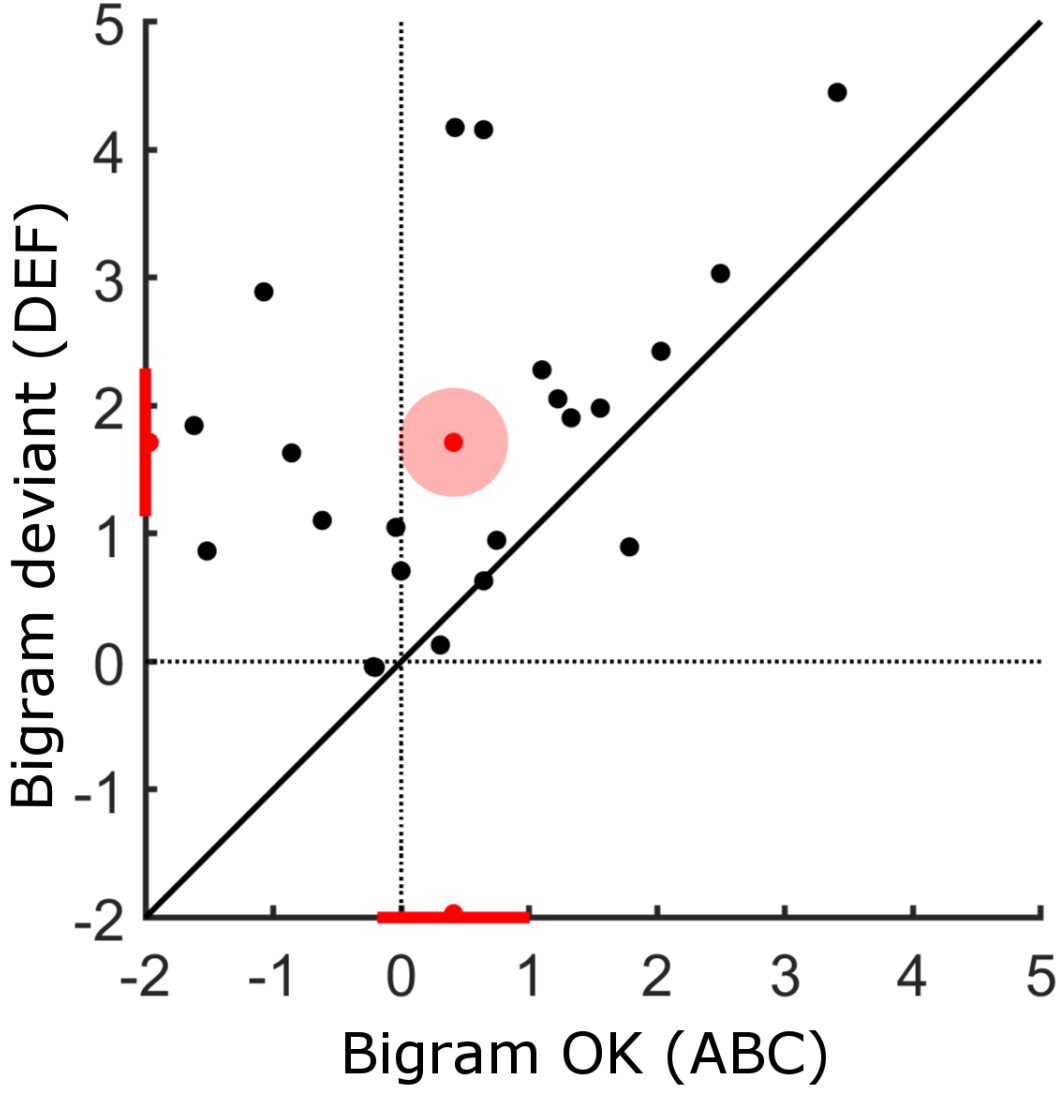
Phantom words in reading, d prime

STIMULUS PRESENTED

RESPONSE GIVEN

	STD	WD	BD
Correct	Correct rejection	Miss	Miss
Mistaken	False alarm	Hit	Hit

dprime (d')

z(Hit) - z(False Alarm)

Phantom words experiment, 6–letter words