Statistical learning and learning to read

Davide Crepaldi

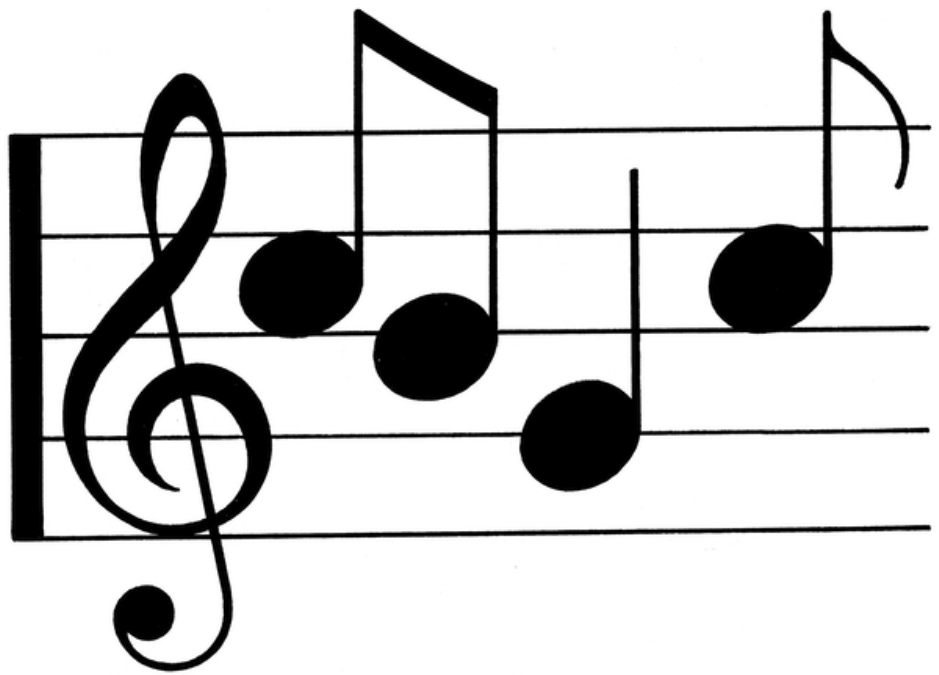
SISSA
Trieste, Italy

davide.crepaldi@sissa.it
Tweets at @CrepaldiDavide

EPS meeting
Symposium in honour of Kate Nation
31 March 2022

Lots of regularities

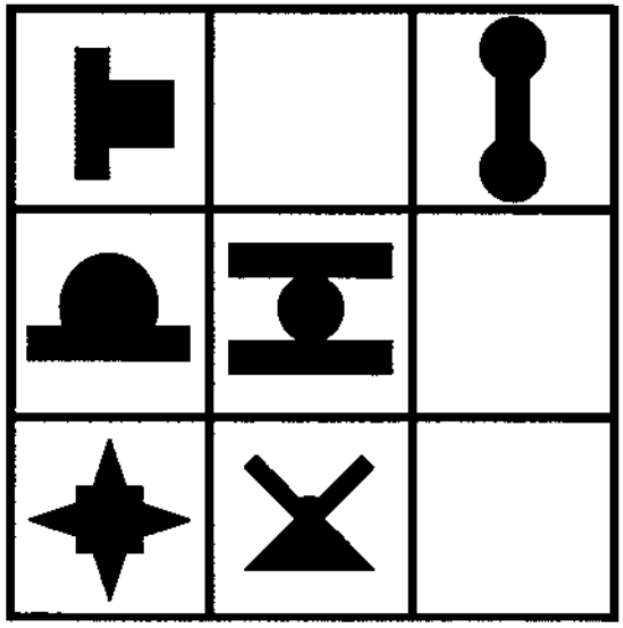
Lots of regularities

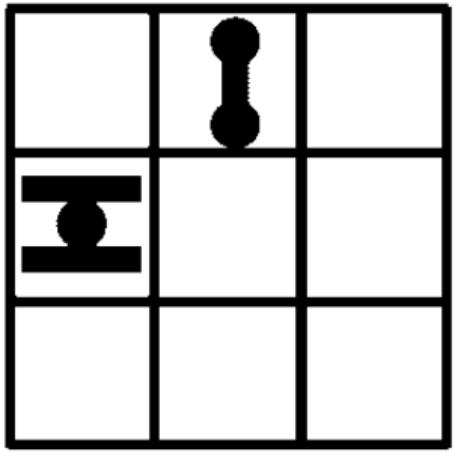


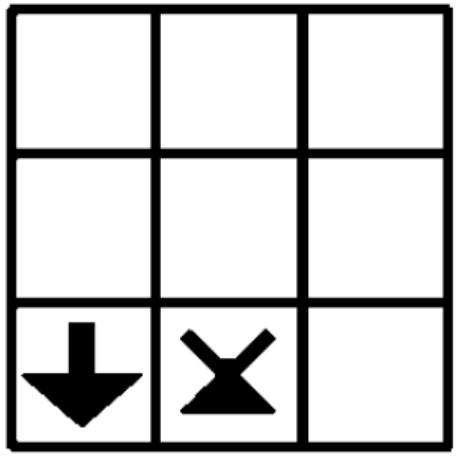
Lots of regularities

Lots of regularities

Statistical learning in visual scenes







(Fiser and Aslin, 2001)

Regularities in the lexicon

- mu vs. cz (**mu**se vs. **cz**ar)

Do we use this information as we process written words?

Not entirely clear

Of course we do

- We distinguish frequent from rare bigrams (e.g., Chetail, 2017)

Nah

- Bigram frequency effects are shaky (e.g., Schmalz and Mulatti, 2017; Owsowitz, 1963; Chetail et al., 2014)

Adults: booo! Children: yeah!

nGram frequency effects in children

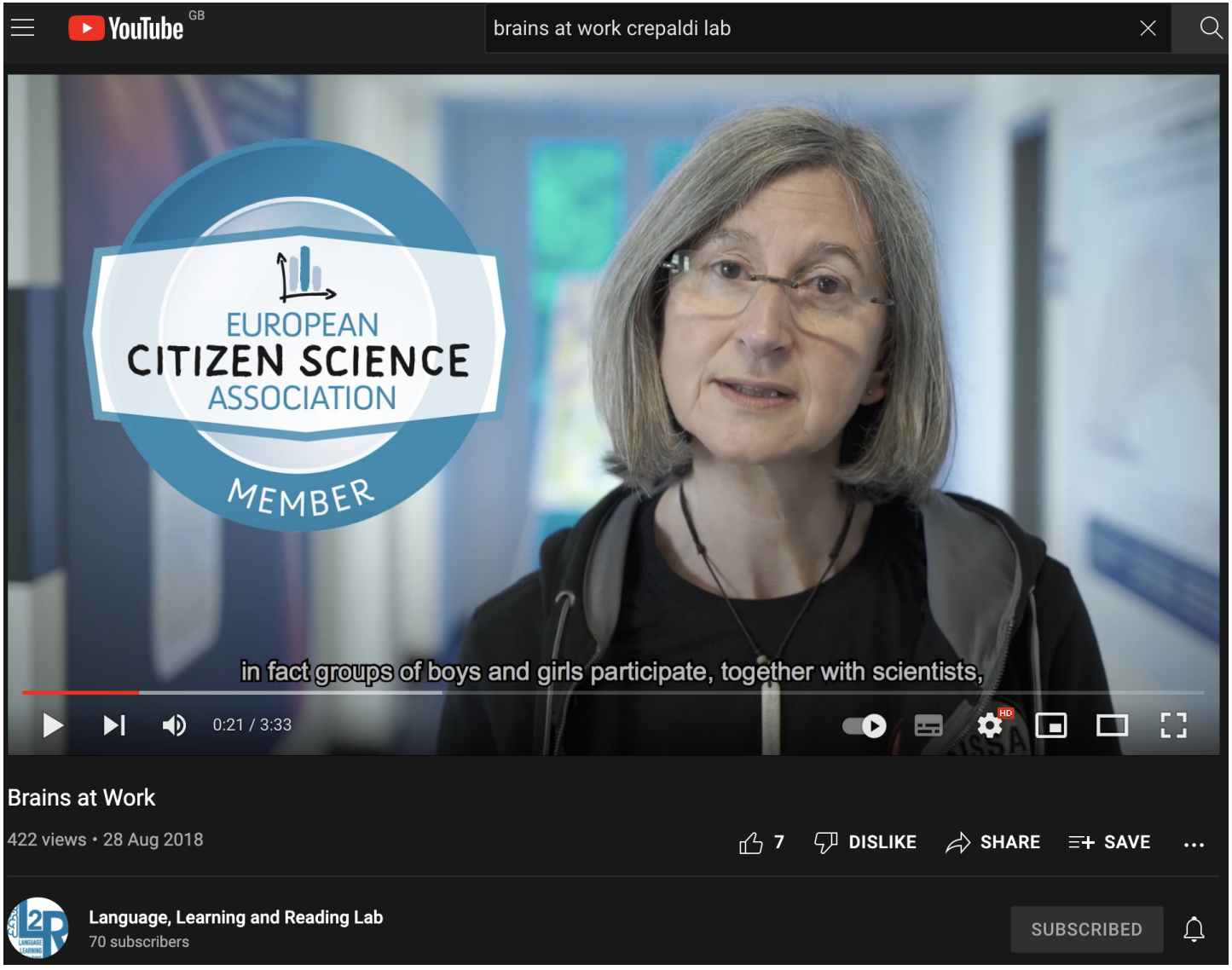
Eye movement and natural reading

- 12 passages, read for comprehension
- Measured eye movements and looked for nGram effects
- Average length: 130.5 (range: 109-170)
- 1566 word tokens, 762 types

	3rd grade	4th grade	5th grade	6th grade	Adults
N	37	20	41	43	33
Age	8.22 (.42)	9.22 (.41)	10.05 (.44)	10.98 (.34)	23.39 (3.32)

- ~200K fixations
- First-of-many fixations (~28K) and gaze durations (~116K)

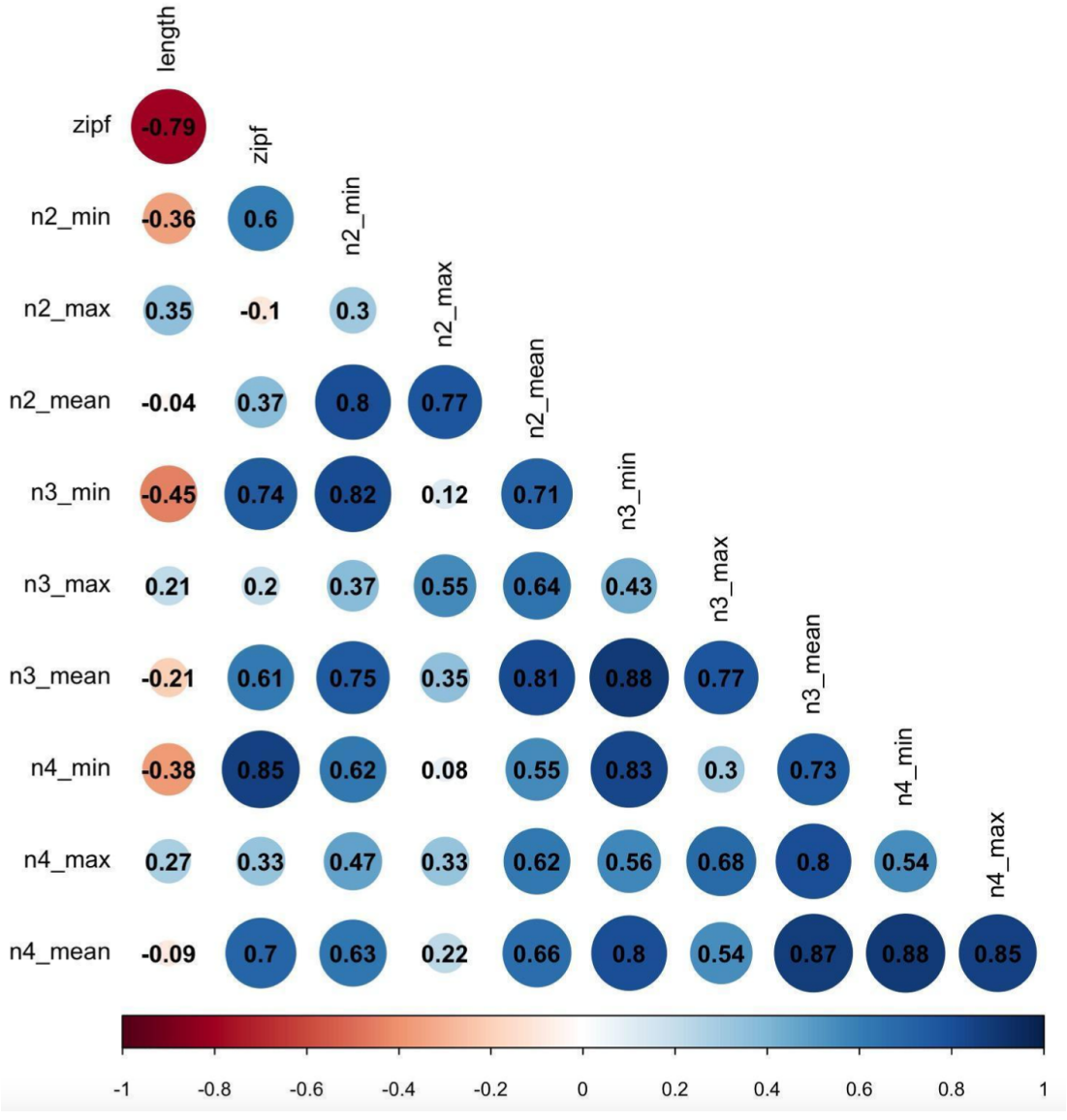
Brains at work



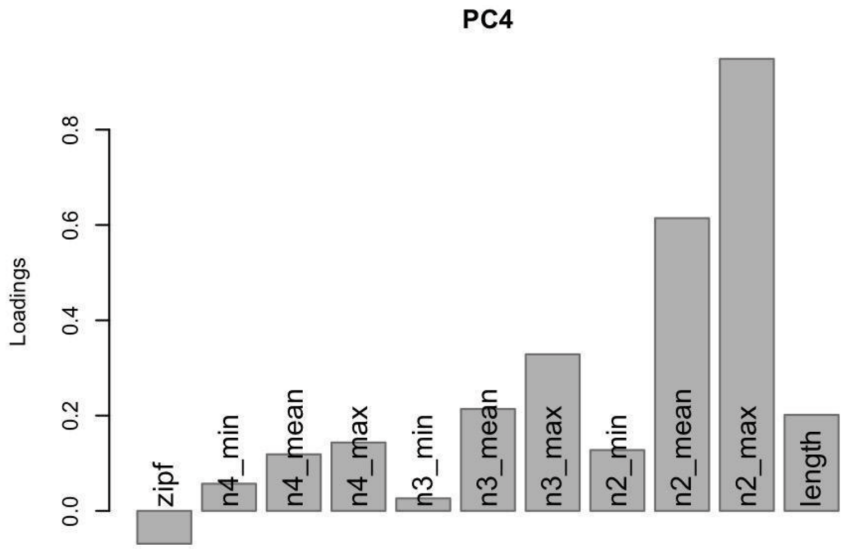
nGram frequency

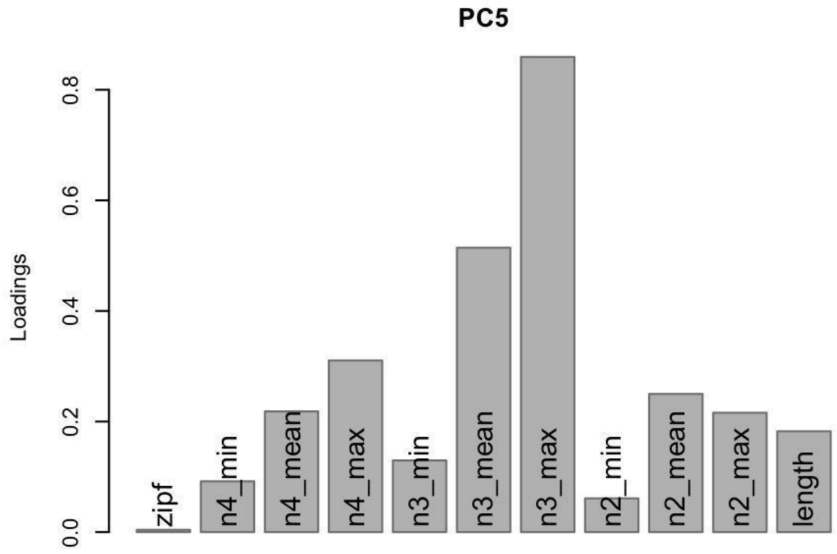
- 2, 3 and 4grams
- Min, max and average

Lots of collinearity. . .

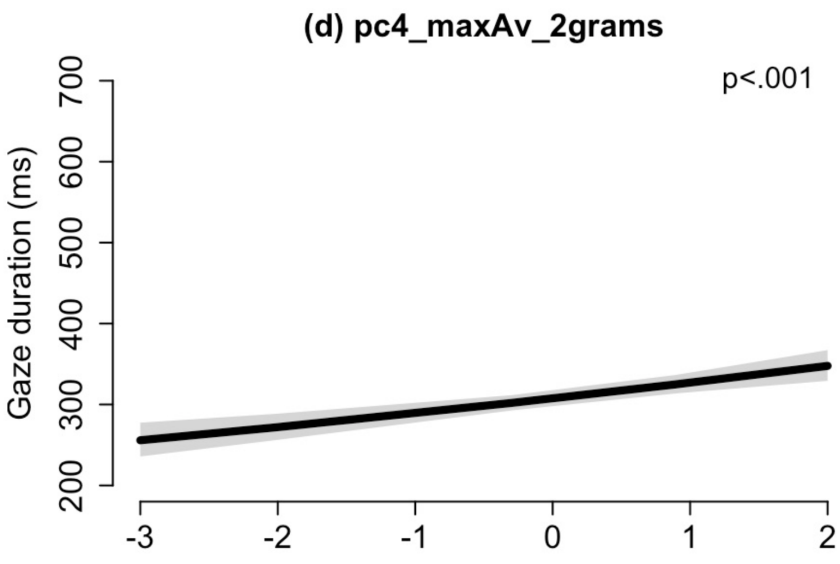


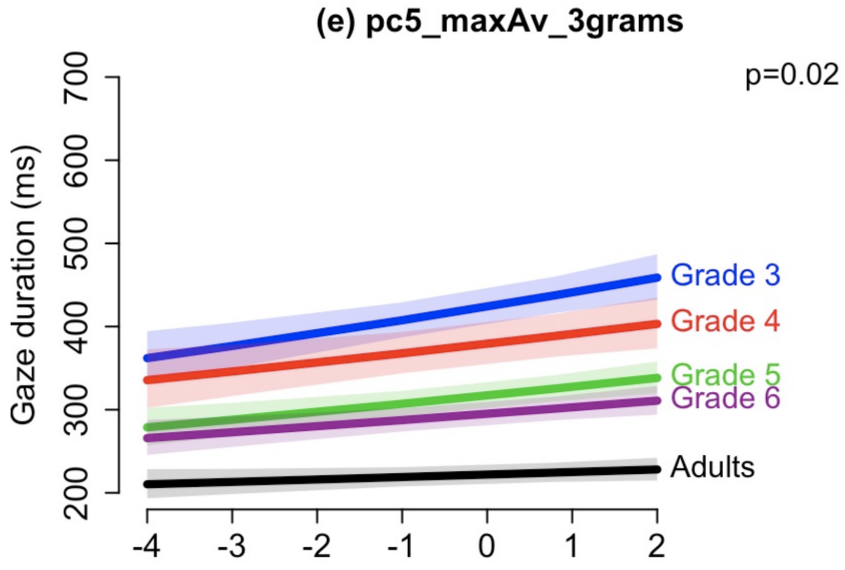
. . . which can be solved



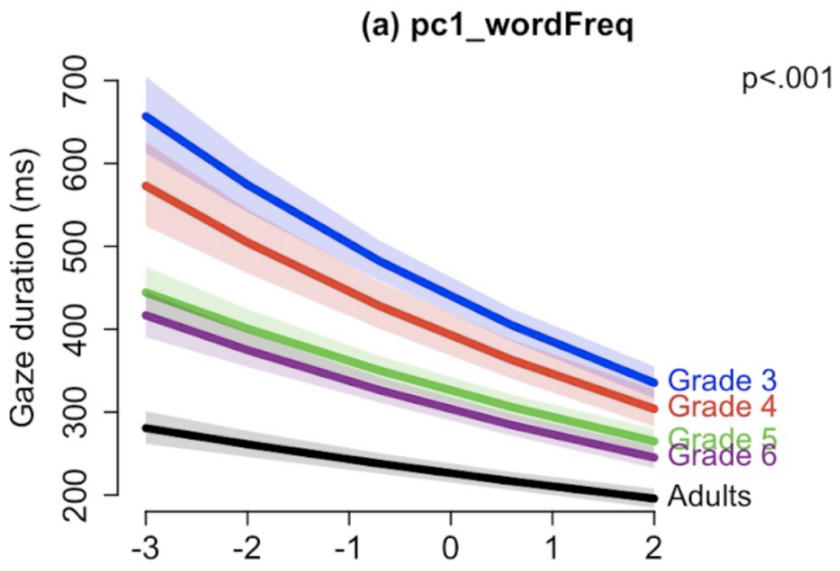


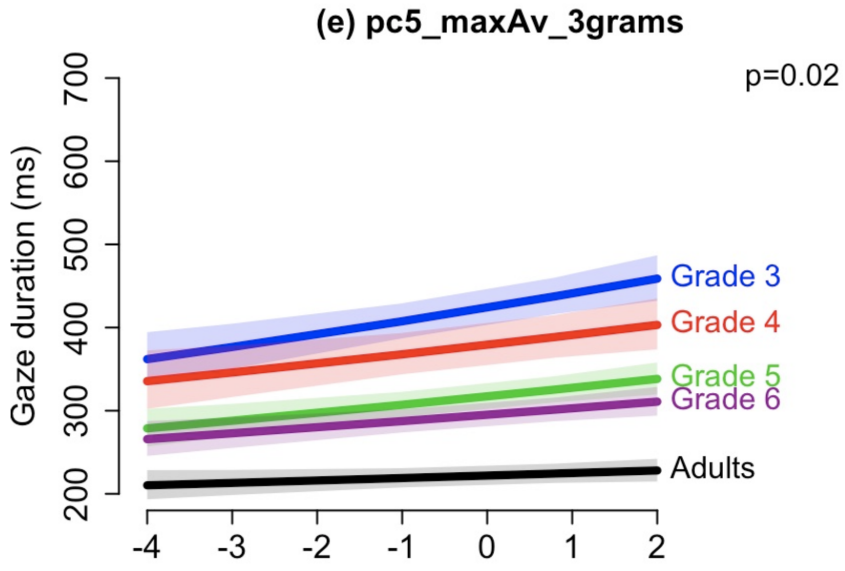
Gaze duration





They're there, but they're small





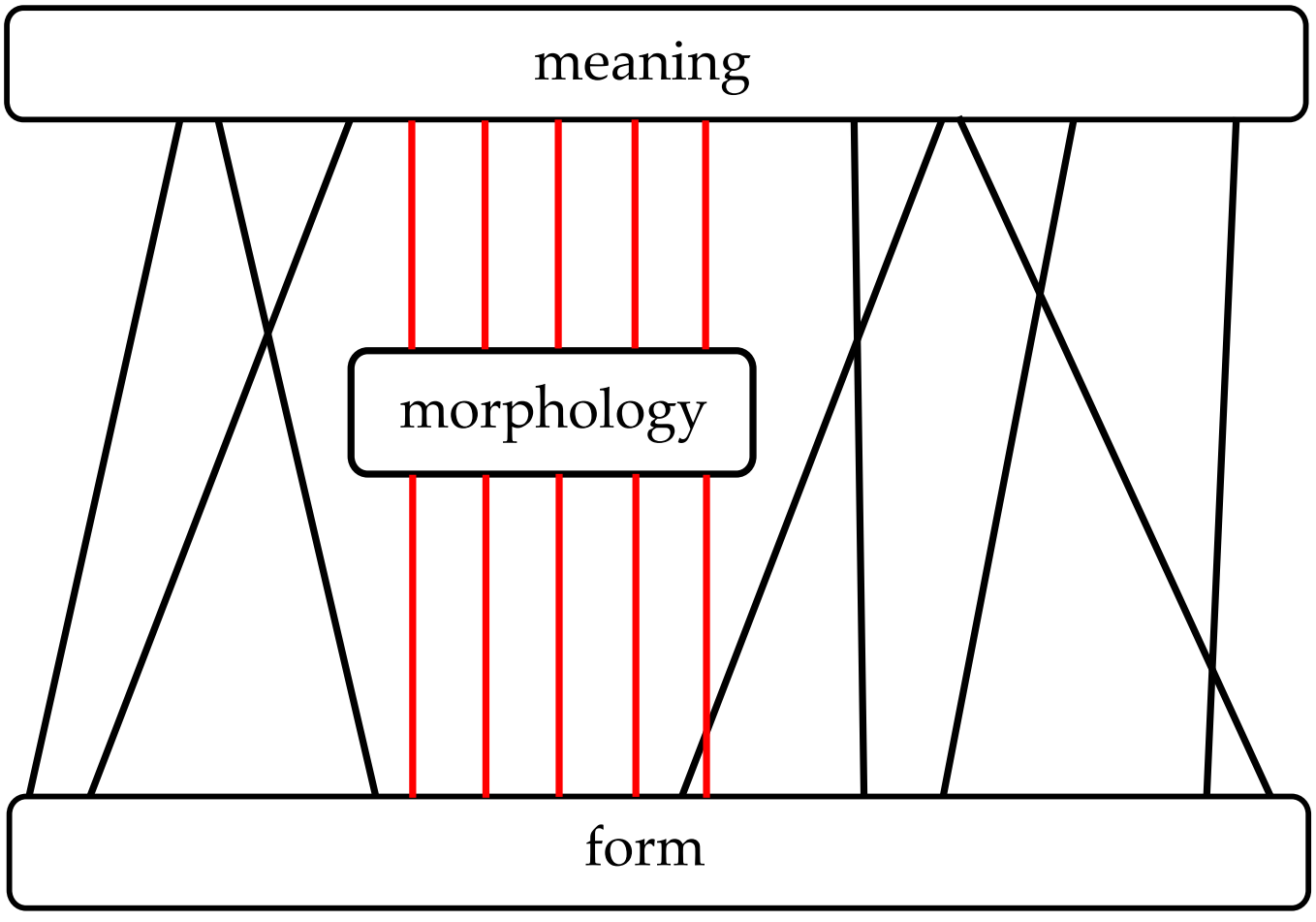
To wrap up

- Children are sensitive to nGram statistics already in Grade 3
- The effects are small though
- Some developmental pattern

- Open question: the role of the spoken language

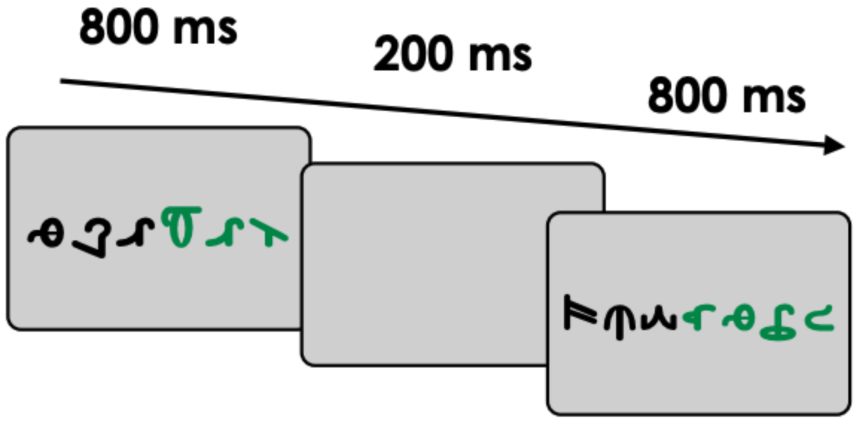
Affix detection based on visual regularity

A breach into language arbitrariness

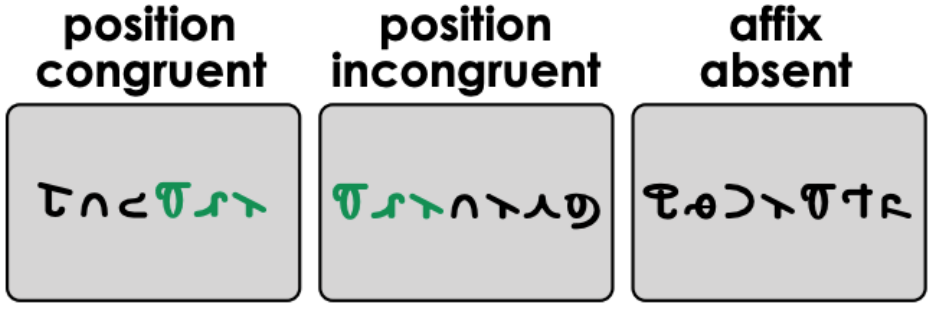


Artificial affixes

Training



Testing

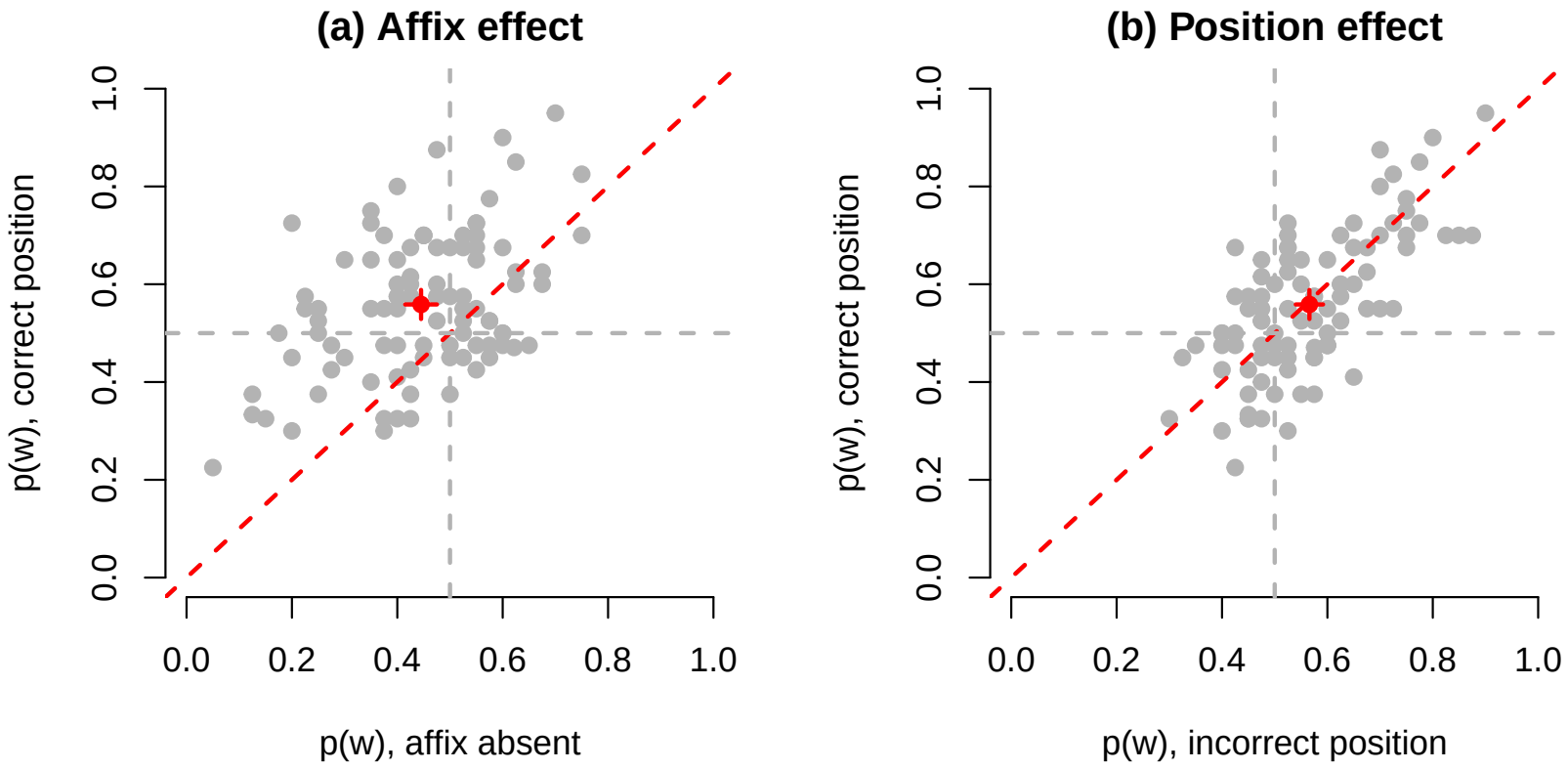


Participants and stimuli

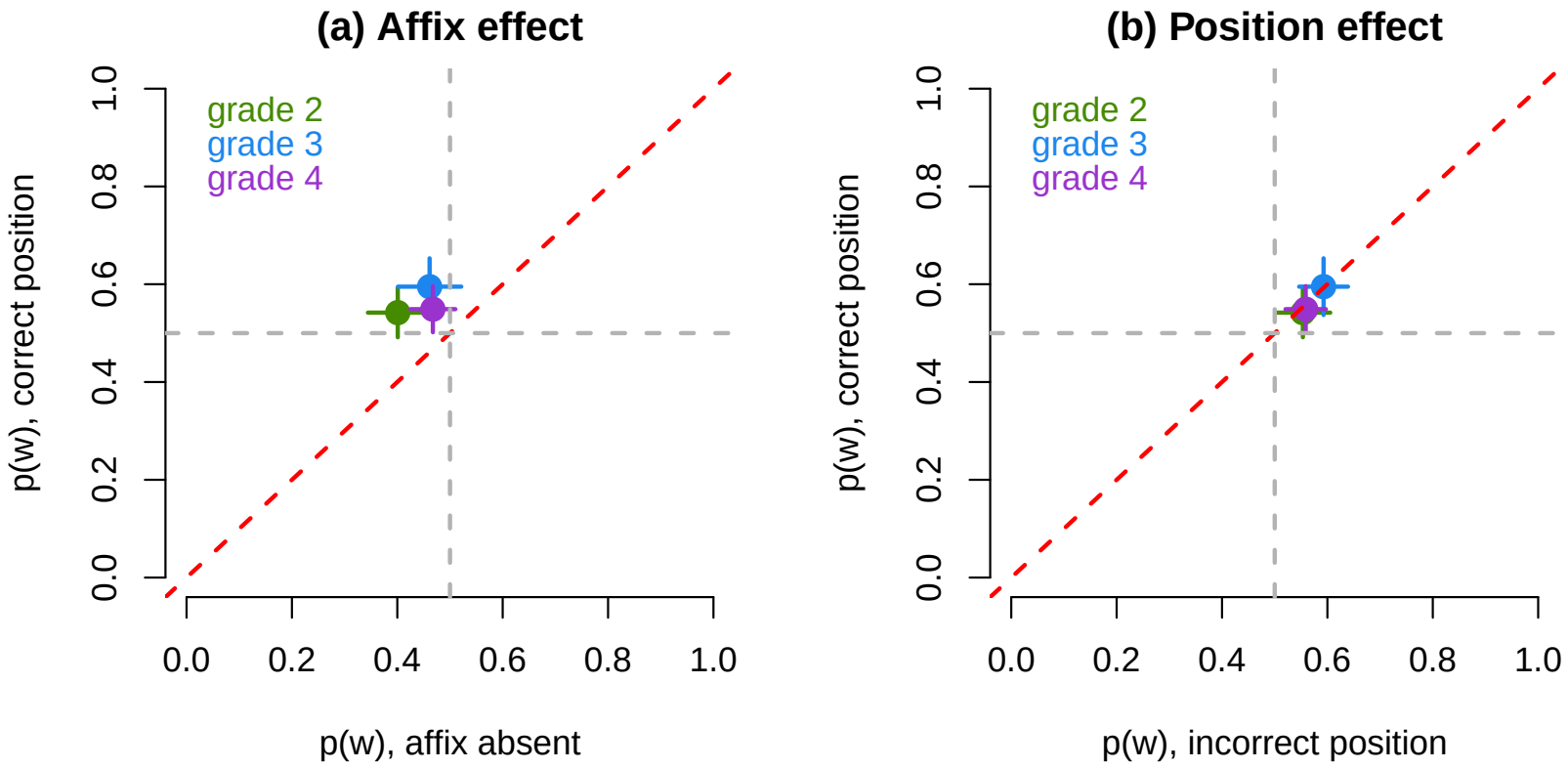
- 5 affix-like chunks, each in 20 "words" (100 training items)
- 120 novel strings for testing, 40 per condition
- Morphological awareness

	1st grade	2nd grade	3rd grade	4th grade	5th grade
N	13	29	24	40	14
Age	6.75 (.55)	7.68 (.29)	8.81 (.44)	9.65 (.39)	10.86 (.35)

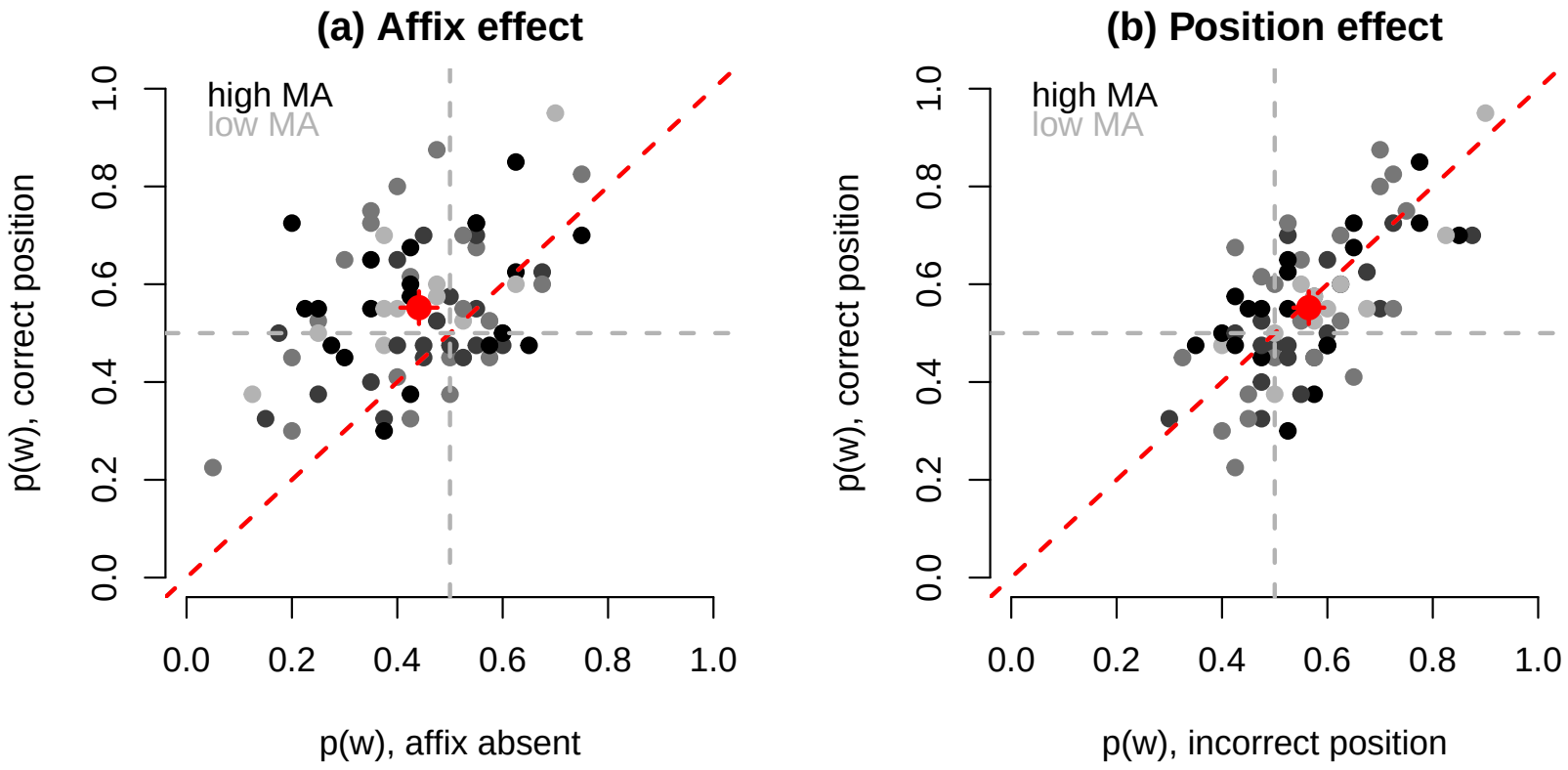
Sensitive to affixes, not to position



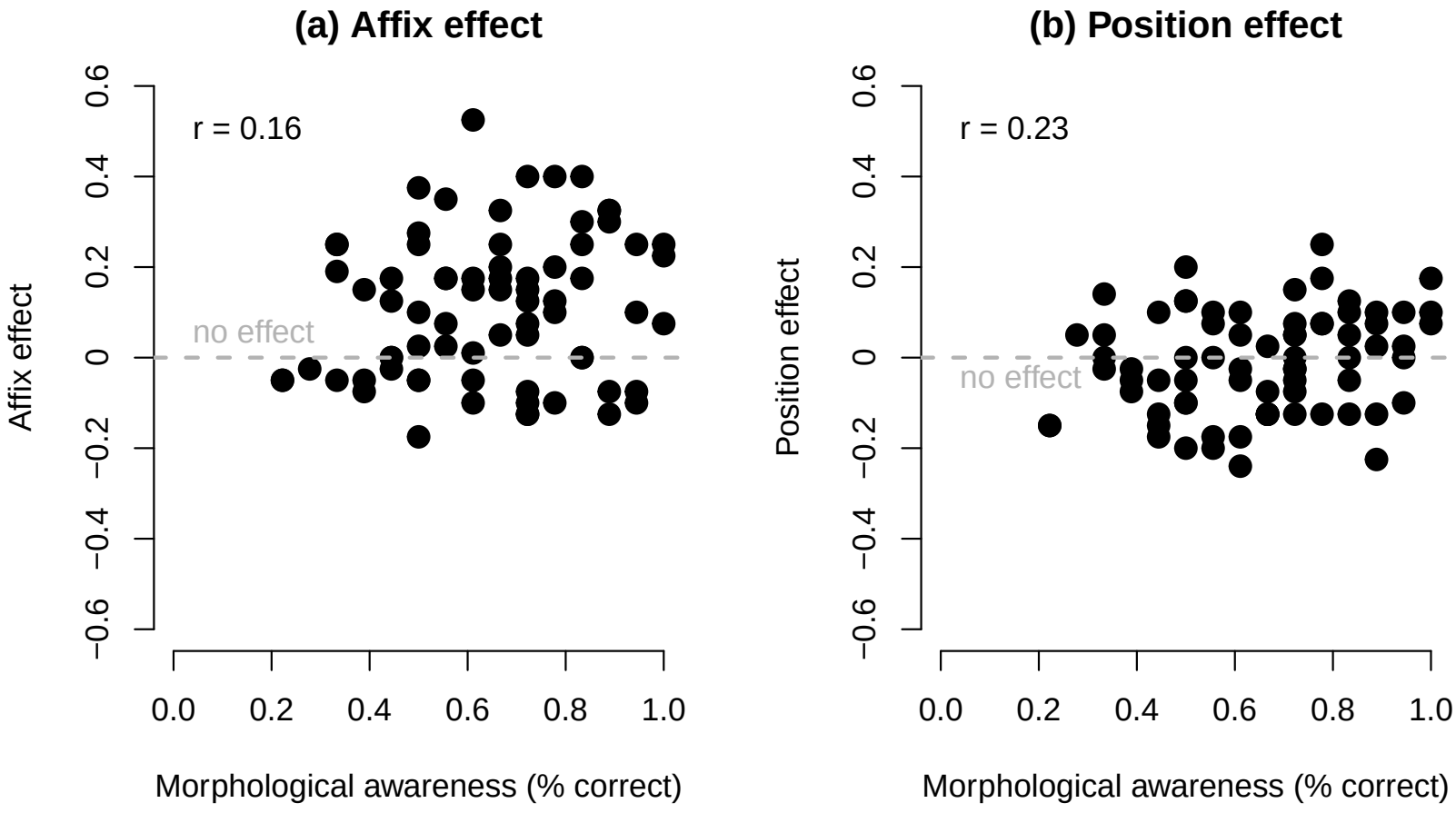
Little developmental pattern



Little (no?) role for morphological awareness



Little (no?) role for morphological awareness



The alien affixes

- Children spontaneously extract visual statistical regularities in strings of novel letters
- This affects lexical judgments
- These skills are already in place in Grade 2
- No evidence for position sensitivity (differently from the adults; Lelonkiewicz et al., 2020)
- Little (no?) role for morphological awareness

To wrap up

- Children do show sensitivity to letter/symbol statistics in strings
- Developing vs. more established representations (children vs. adults)

- What's the mechanism (e.g., phonology, meaning)?
- Causal effect in the **actual** learning to read?

Acknowledgments and links

- Jarosław Lelonkiewicz
- Maria Ktori
- Valentina Pescuma
- Jon Carr

- Brains at work video

- The lab website

Statistical learning and learning to read

Davide Crepaldi

SISSA
Trieste, Italy

davide.crepaldi@sissa.it
Tweets at @CrepaldiDavide

EPS meeting
Symposium in honour of Kate Nation
31 March 2022

References I

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

Chetail, F., Balota, D., Treiman, R., and Content, A. (2014). What can megastudies tell us about the orthographic structure of english words? *Quarterly Journal of Experimental Psychology*, 68:1519–1540.

Crepaldi, D., Rastle, K., and Davis, C. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory and Cognition*, 38(3):312–321.

Fiser, J. and Aslin, R. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6):499–504.

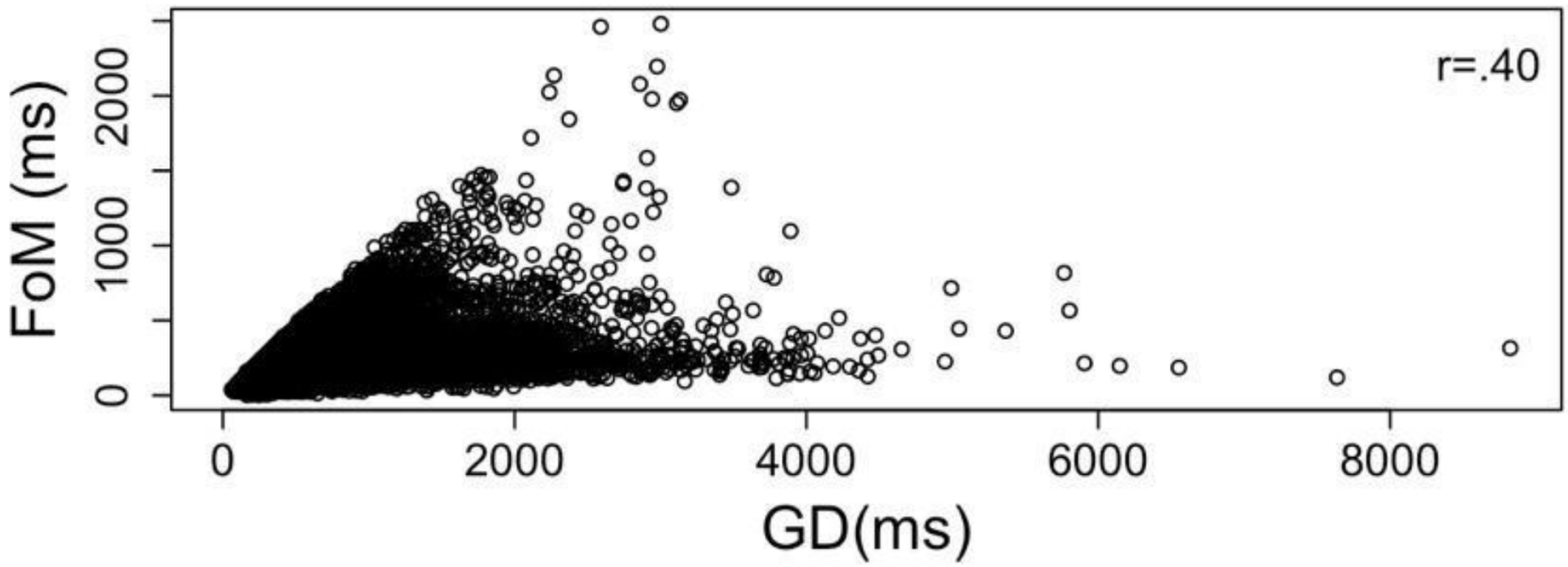
Lelonkiewicz, J., Ktori, M., and Crepaldi, D. (2020). Morphemes as letter chunks: Discovering affixes through visual regularities. *Journal of Memory and Language*, page 104152.

Owsowitz, S. E. (1963). *The effects of word familiarity and letter structure familiarity on the perception of words*. Rand Corporation Publications.

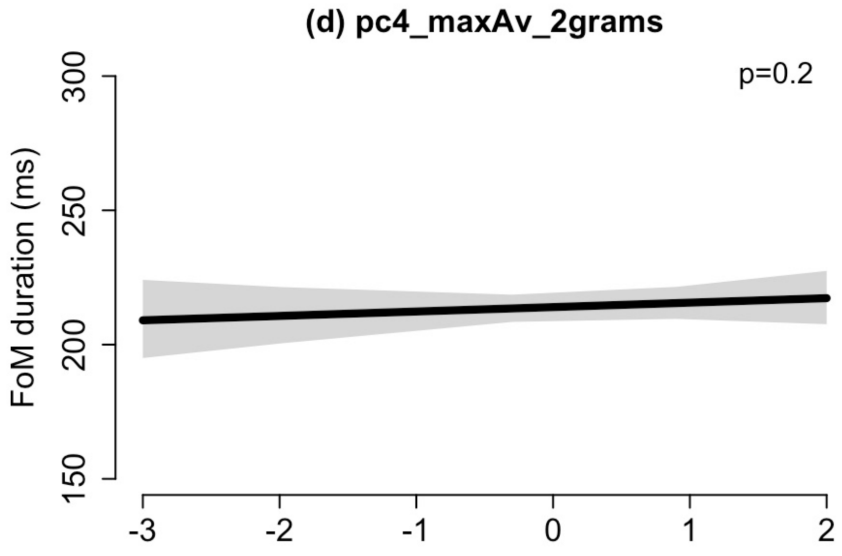
Rastle, K., Davis, M., and New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6):1090–1098.

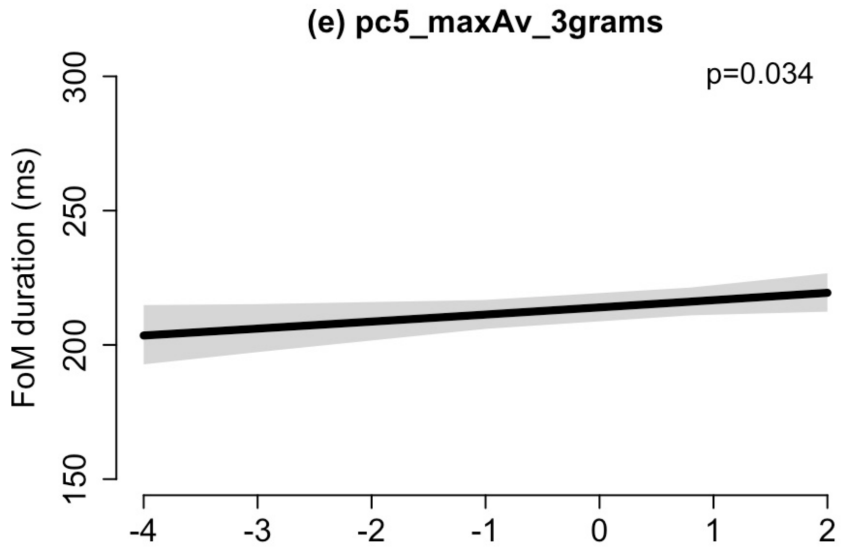
Schmalz, X. and Mulatti, C. (2017). Busting a myth with the bayes factor: Effects of letter bigram frequency in visual lexical decision do not reflect reading processes. *The Mental Lexicon*, 12:263–282. Retrieved from osf.io/3ybwd.

FoM and gaze duration

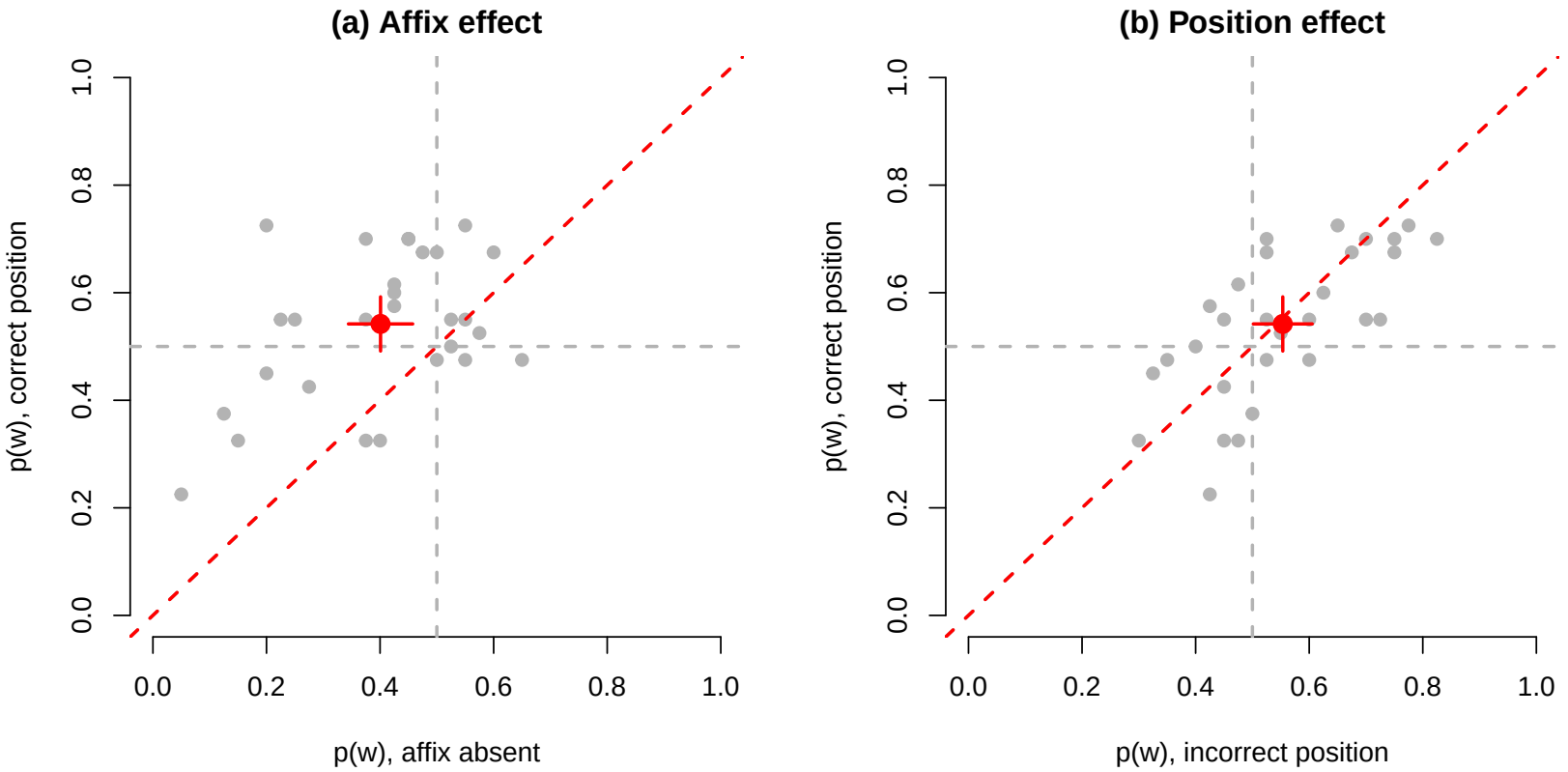


First-of-many fixations

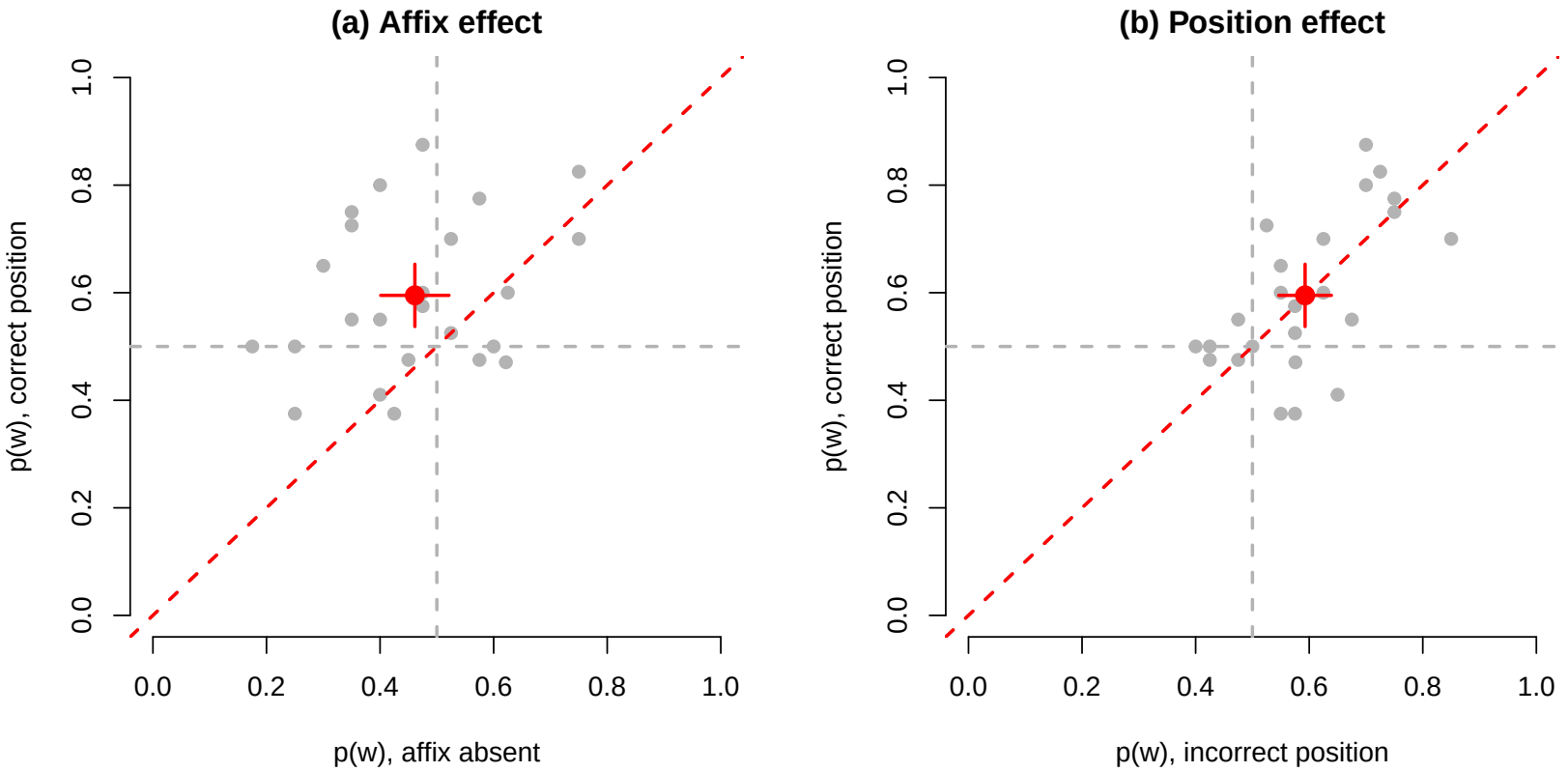




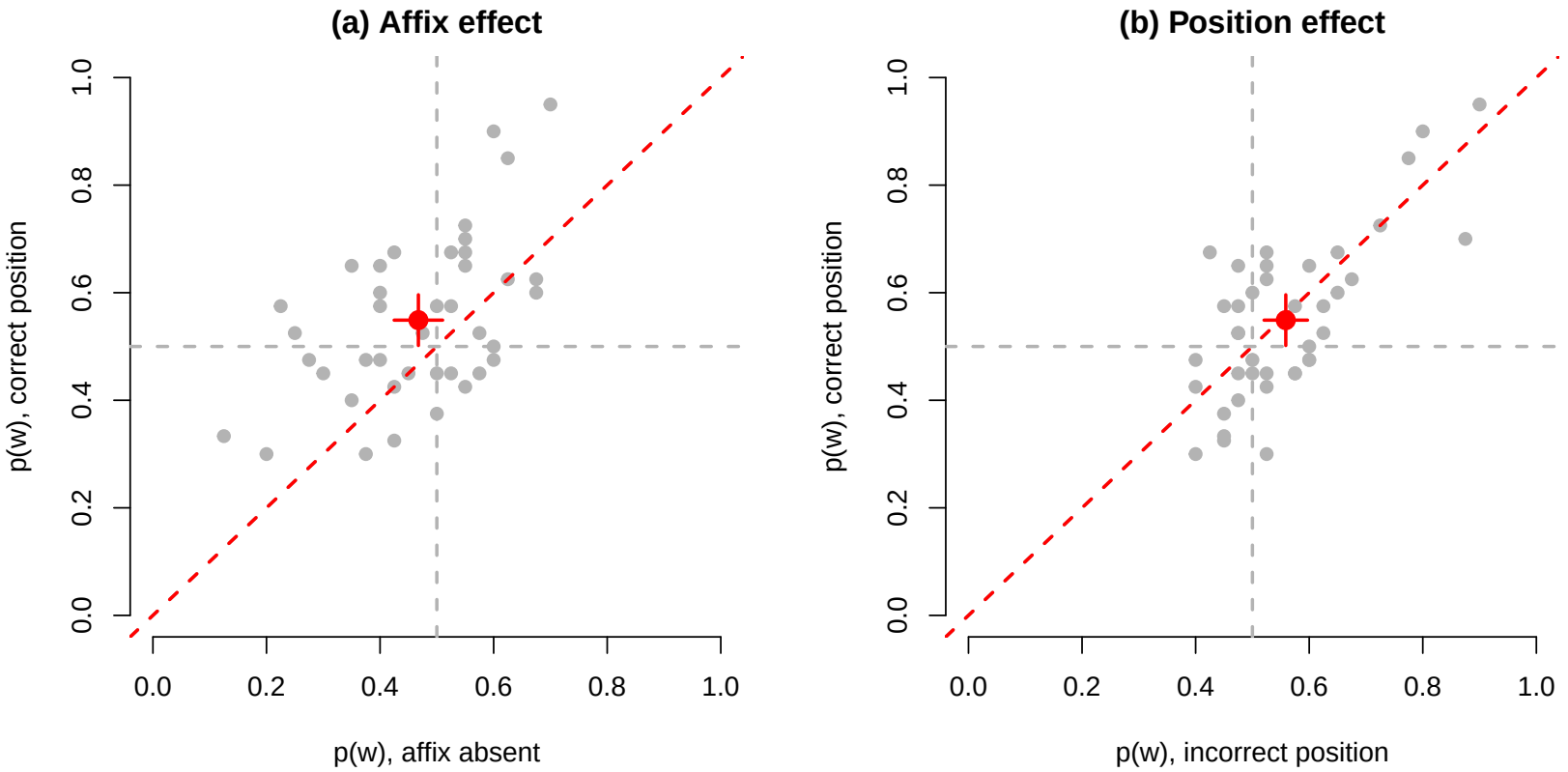
Affixes make strings words, grade 2



Affixes make strings words, grade 3

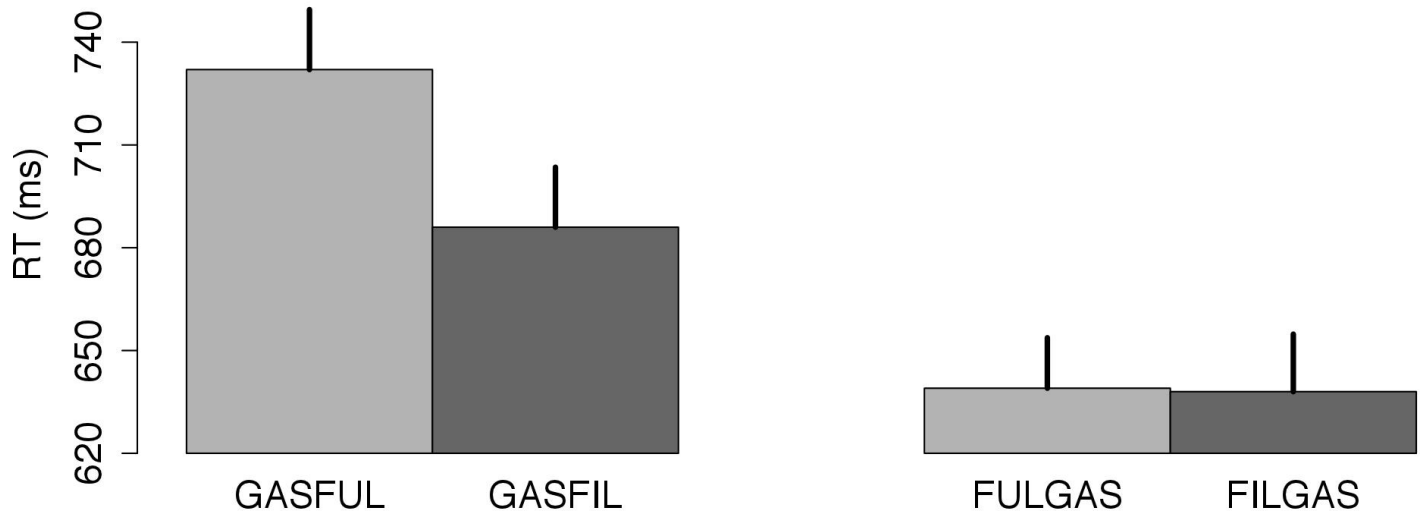


Affixes make strings words, grade 4



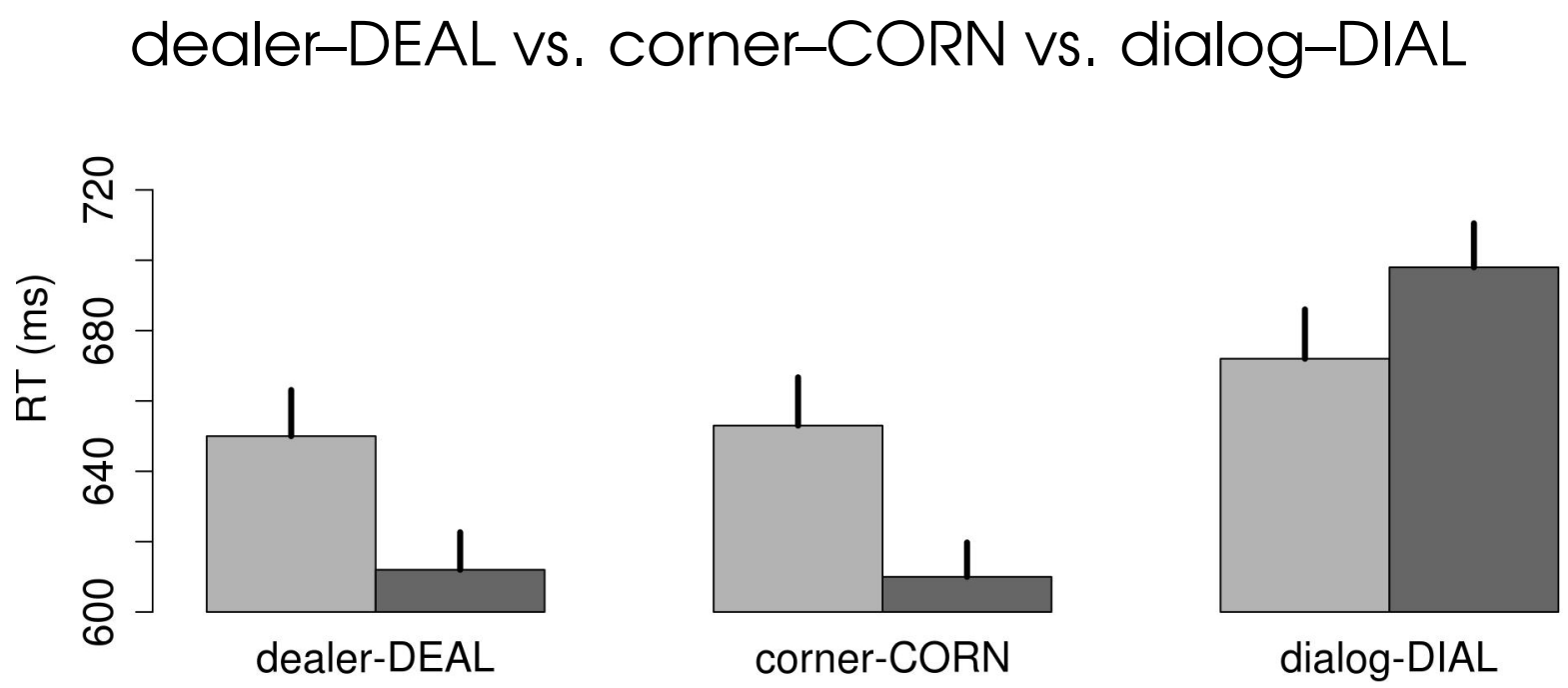
Blind to suffixes?

- (GASFUL vs. GASFIL) vs. (FULGAS vs. FILGAS)



(Crepaldi et al., 2010)

Corners that corn



(Rastle et al., 2004)