The Reading Brain as a Statistical Learning Machine

Davide Crepaldi

International School for Advanced Studies (SISSA), Trieste

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

Heriot–Watt University, Edinburgh, 26 September 2018

Reading is a human wonder

We’re extremely good readers . . .

- 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 ~300 words per minute (e.g., Pelli et al., 2007)

. . . with no genetic endowment

- Written language isn't observed universally
- Literacy isn't acquired spontaneously

Word morphology

Arbitrariness

- elephant
- table
- heat
- drum

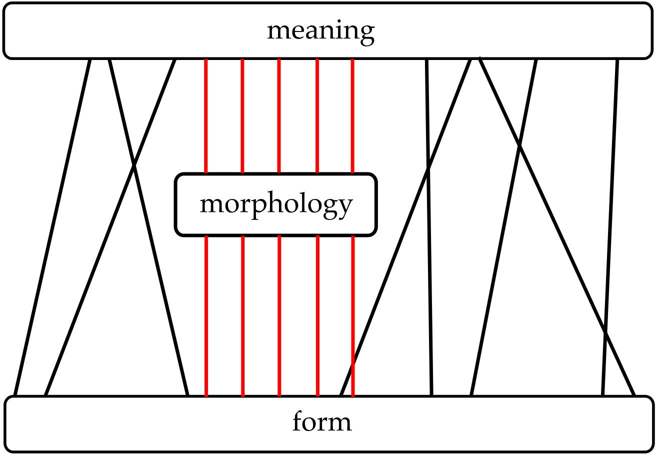
Arbitrariness. Really?

- elephant
- table
- heat
- drum

- preheat
- juicer

- bioweapon
- guesstimate
- untweet (?)

A break into arbitrariness



The core idea

Morphology creates probabilistic regularities in language form and in form–to–meaning mapping. The brain codes for these regularities and uses them during processing.

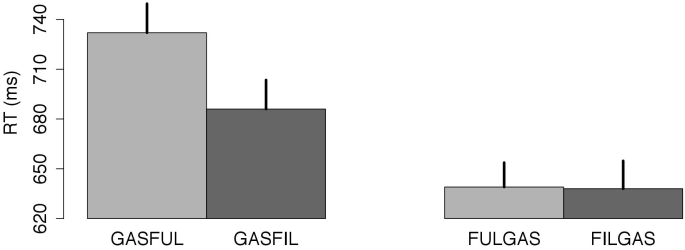
Positional constraints in morpheme perception

Morpheme positional constraints

- KINDNESS and NESSKIND
- PREHEAT and HEATPRE
- CATWALK and WILDCAT
- OVERHANG and HANGOVER

Blind to suffixes?

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

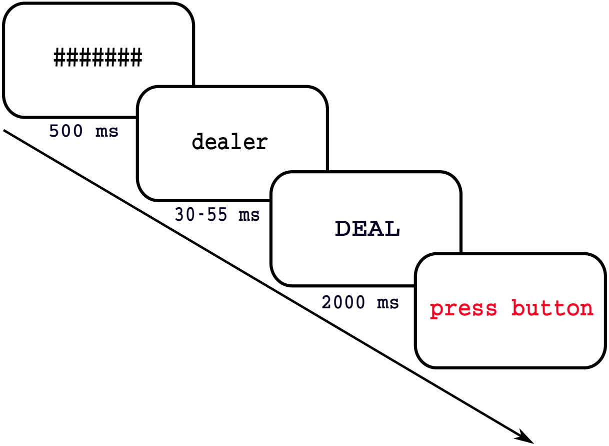


(Crepaldi et al., 2010)

Blind to prefixes?

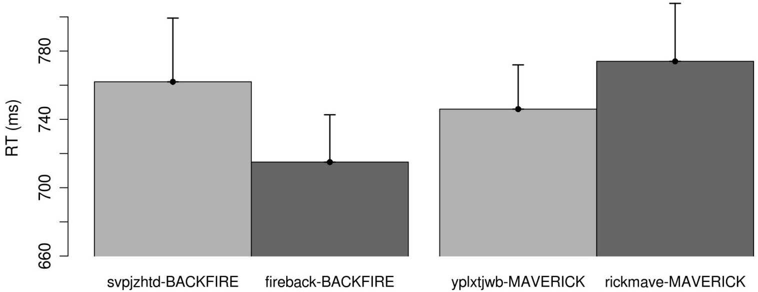
- (PREHOSE vs. PLEHOSE) vs. (HOSEPRE vs. HOSEPLE)

Stems everywhere?



Stems everywhere?

- (fishgold–GOLDFISH vs. kacnvrqw–GOLDFISH) vs. (tonebari–BARITONE vs. suyzchmw–BARITONE)



(Crepaldi et al., 2013)

Orthography–to–Semantic Consistency (OSC)

Form as a cue to meaning

CORN

- Get all words that start with CORN
- Take their semantic representations
- Compute their similarity
- Take the mean

Orthography–Semantic Transparency (OSC)

- The internal consistency of the ‘form’ family in terms of meaning
- How similar in meaning are words similar in form
- How good of a cue to meaning is form

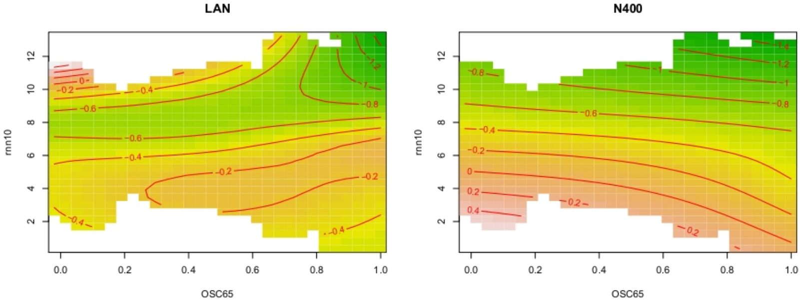
OSC gets unique variance in lexical decision times

Table 6. *Results of the regression analysis on the lexical decision latencies extracted from the BLP for a large set of random words*

	Estimate	*Std error*	t *value*	p *value*
Intercept	6.5922	.0109	602.89	.0001
Word frequency	−0.0308	.0009	33.41	.0001
Word FS	−0.0041	.0021	1.97	.0495
Word length	0.0035	.0013	2.74	.0061
OSC	−0.0254	.0066	3.84	.0002

(Marelli et al., 2014)

OSC also explains ERPs



(Amenta et al., in prep.)

The core idea

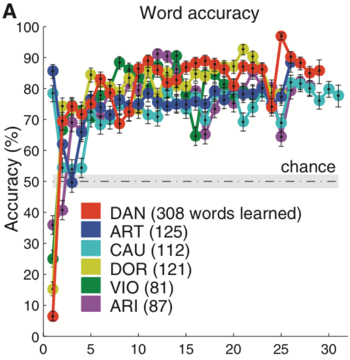
Morphology creates probabilistic regularities in language form, and in form–to–meaning mapping. The brain codes for these regularities and uses them during processing.

The core idea

Language shows probabilistic regularities in its form, and in form–to–meaning mapping. The brain codes for these regularities and uses them during processing.

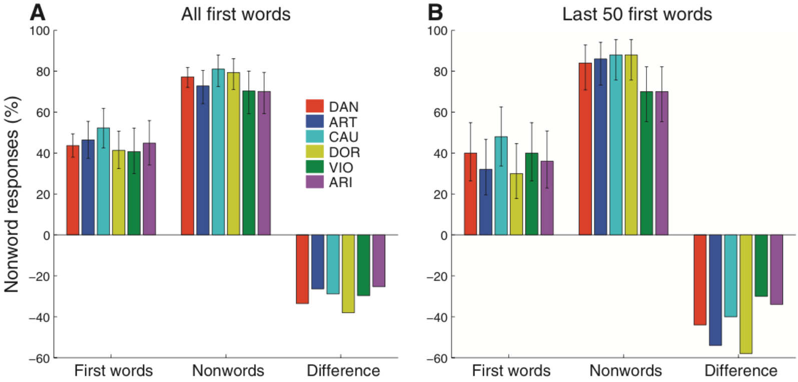
Orthography in Baboons

Baboons learn words

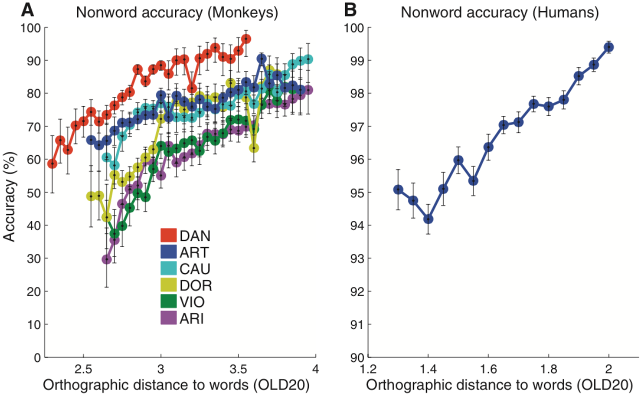


(Grainger et al., 2012)

Baboons extract knowledge about letter stats



Baboons extract knowledge about letter stats



The lesson from Baboons

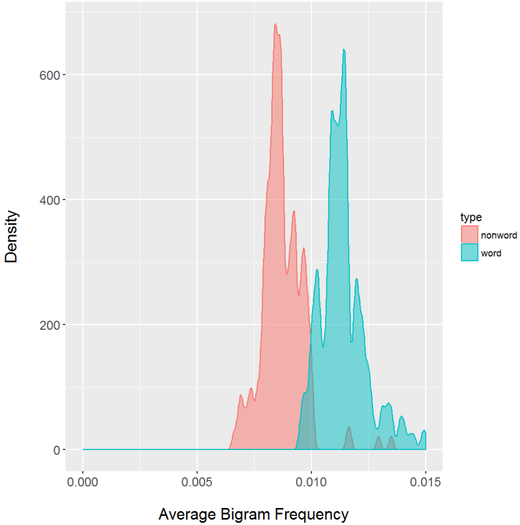
- We don't need language to do visual word identification
- Visual word learning proceeds through letter stats (perhaps)

Novel word learning in humans

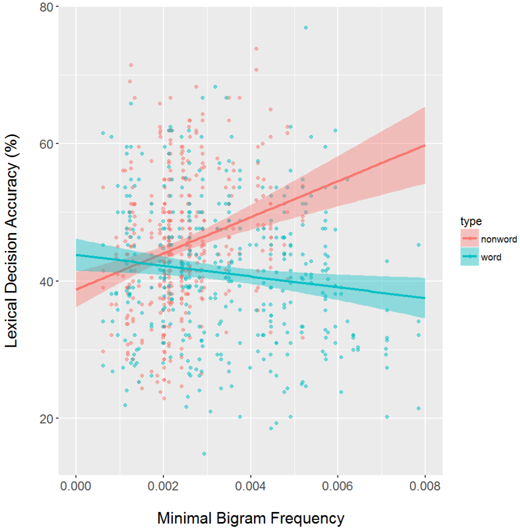
A new lexicon

- 200 novel words (e.g., mefoal), 6 to 9 letters long
- Each novel word presented 3 times, interspersed with 600 non–words (e.g., paltoon)
- Lexical decision with feedback

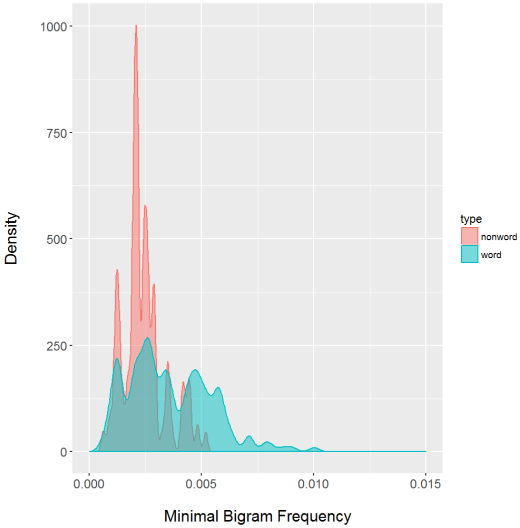
Letter stats distinguish words and non-words in the novel language



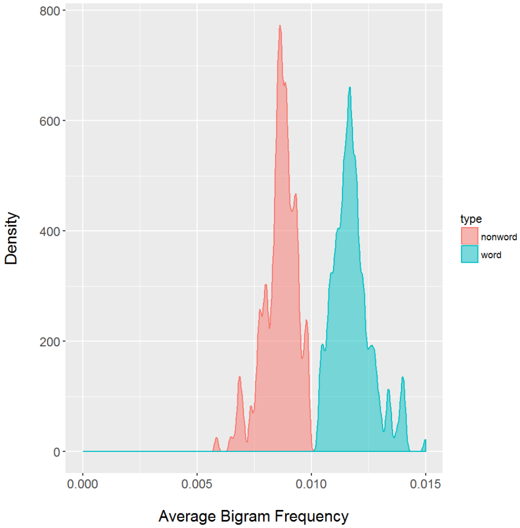
Minimal bigram frequency rules



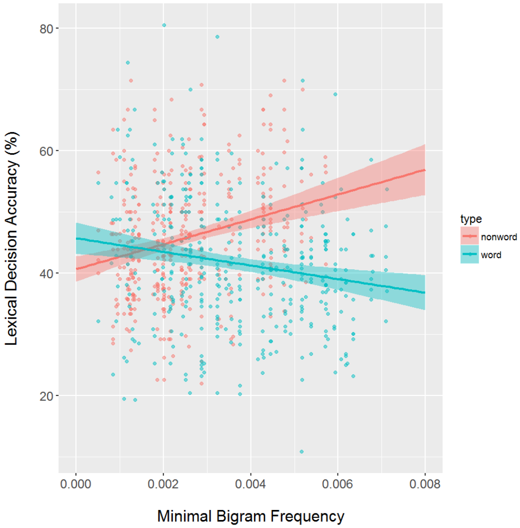
Not a useful cue



A stronger cue



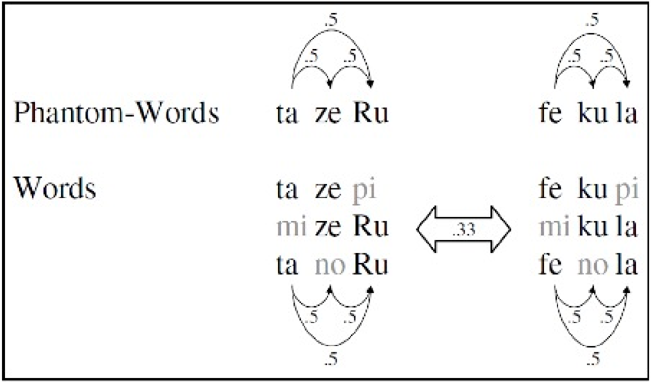
Still, minimal bigram frequency is what matters



The lesson from humans

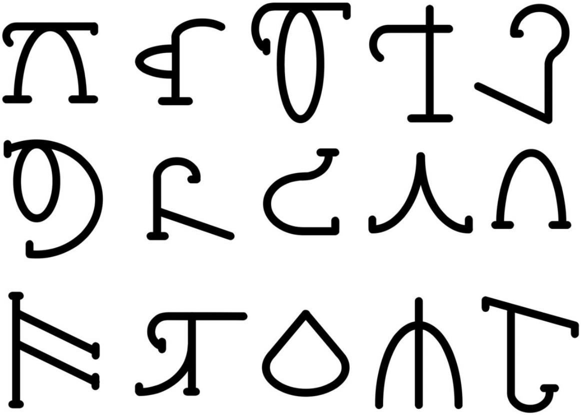
- We do code for letter stats
- We don't seem to figure out the informative cue in a novel lexicon though, we just go for minimal bigram frequency
- Unfamiliar script?

Phantom words



(Endress and Mehler, 2009)

Pseudofonts



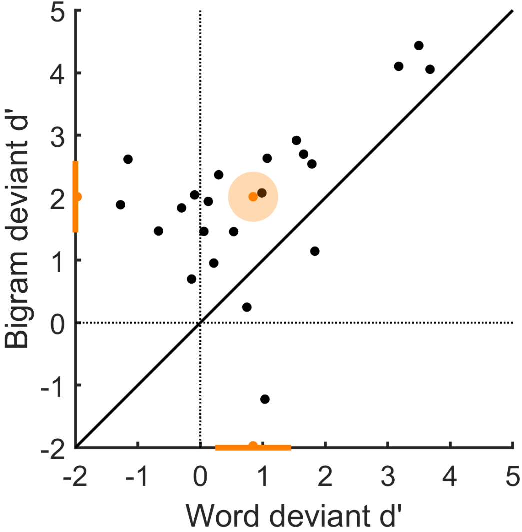
(Vidal et al., 2017)

Phantom words in reading

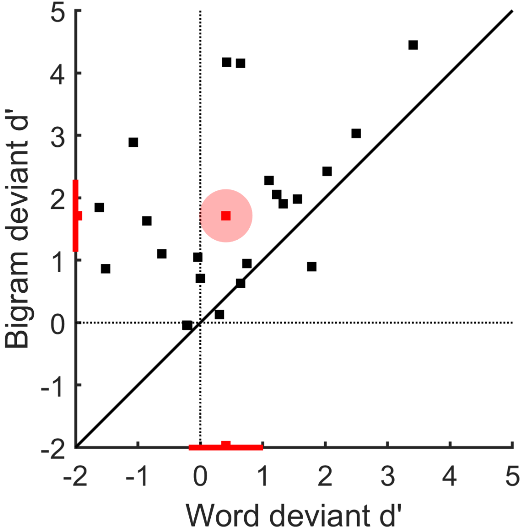
... S S S S D S S S D S S S S D S S ...

	T E P	F U P	
STD	M E R	M U L	90% (15%)
	T O R	F O L	5.55% - 5%
Word Dev	T E R		5% 6.66%
Bigram Dev	M O P		5% 1.66%

Phantom words in reading



Longer words, same story



The lesson from phantom words

- We code for bigrams as we learn novel words
- No strong commitment to bigrams, we didn't try much else—the core point is that we capture sub–lexical stats in a **word** learning task

The core idea

Language shows probabilistic regularities in its form, and in form–to–meaning mapping. The brain codes for these regularities and uses them during processing.

A new approach to reading

- Scripts can be seen as fully–fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning

A new approach to reading

- Scripts can be seen as fully–fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning

Acknowledgments

- The Ministry of Education, University and Research in Italy; the ESRC, the British Academy and the Leverhulm Trust in the UK.
- The ERC Starting Grant 679010 (StatLearn)
- Jarek Lelonkiewicz and Yamil Vidal Dos Santos at SISSA; Simona Amenta in Gent; Marco Marelli at Milano Bicocca; Kathy Rastle at Royal Holloway; Colin Davis at Bristol.

The Reading Brain as a Statistical Learning Machine

Davide Crepaldi

International School for Advanced Studies (SISSA), Trieste

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

Heriot–Watt University, Edinburgh, 26 September 2018

References I

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.

Crepaldi, D., Rastle, K., Davis, C. J., and Lupker, S. J. (2013). Seeing stems everywhere: Position-independent identification of stem morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, 39:510–525.

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., Colé, P., and Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, 30:370–384.

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

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., and Xu, F. (1992). *Overregularization in language acquisition.* Monograph of the Society for Research in Child Development.

References II

Marelli, M., Amenta, S., and Crepaldi, D. (2014). Semantic transparency in free stems: The effect of Orthography–Semantics Consistency on word recognition. *Quarterly Journal of Experimental Psychology*, pages 1–13.

Pelli, Denis G. anf Tillman, K. A., Freeman, J., Su, M., Berger, T. G., and Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, 7(2):20:1?36.

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.

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

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7:263–272.

Taft, M. and Forster, K. I. (1975). Lexical stroage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 15:638–647.

Tamminen, J., Davis, M. H., and Rastle, K. (2015). From specific examples to general knowledge in language learning. *Cognitive Psychology*, (in press).

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.