PSEUDO-LETTER CHUNKING IN NOVEL WORDS THROUGH PROBABILISTIC INFORMATION

Jarosław R. Lelonkiewicz, Maria Ktori & Davide Crepaldi
International School for Advanced Studies (SISSA), Trieste, Italy
jlelonki@sissa.it

BACKGROUND

Morphemes are frequent co-occurring letter chunks with semantic or syntactic properties (e.g., the suffix -er in *dealer* and *player* denotes an agent). During visual word processing, readers recognise morphemes and code for their typical position within words [1,2]. But how do we construct morpheme representations? Drawing from recent evidence from psycholinguistics and statistical learning [3], we test whether morpheme representations are based on a letter-chunking mechanism that utilises probabilistic information in the visual input.

RESEARCH QUESTION

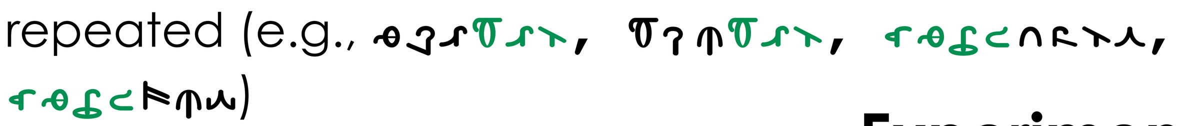
To what extent is letter chunking due to a general ability of the visual system to extract statistical regularities?

METHODS

LEARNING PARADIGM

1. EXPOSURE PHASE

- Passive viewing of 200 pseudo-letter [4] strings
- Strings made of affix-like & stem-like chunks defined by their statistics of occurrence in stimulus set
- Affix-like chunks repeated 20 but stem-like chunks never repeated (e.g.,

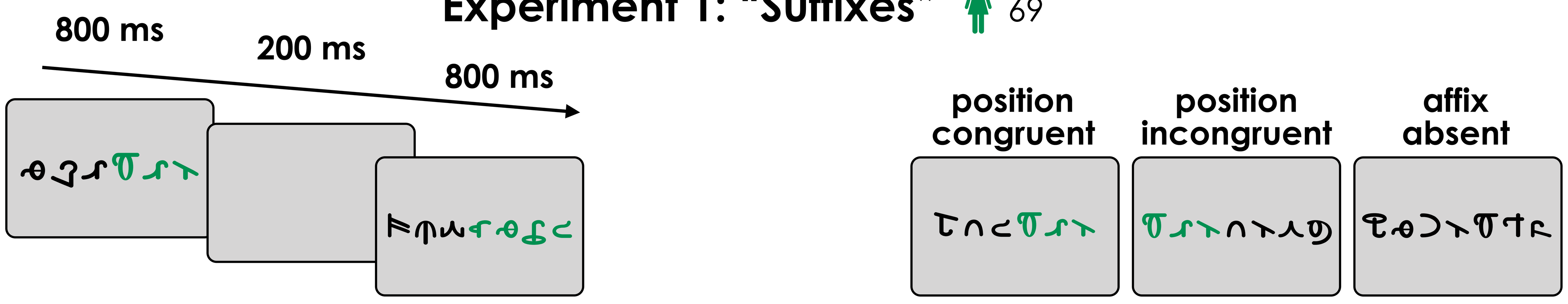


)

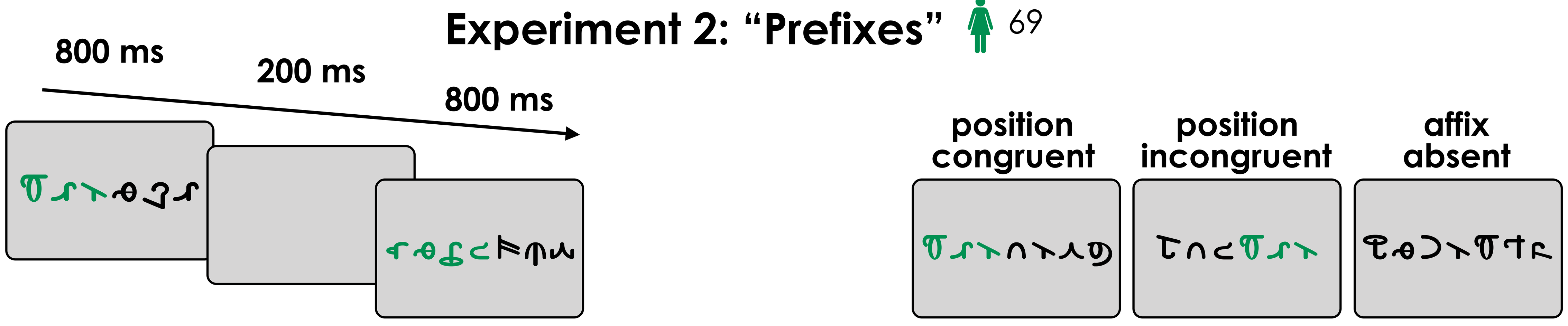
2. JUDGMENT TASK

- Does the string belong to the same language seen in the exposure phase?
- Yes/No keypress
- 300 novel pseudo-letter strings
- Stimulus duration: response terminated

Experiment 1: "Suffixes" 69



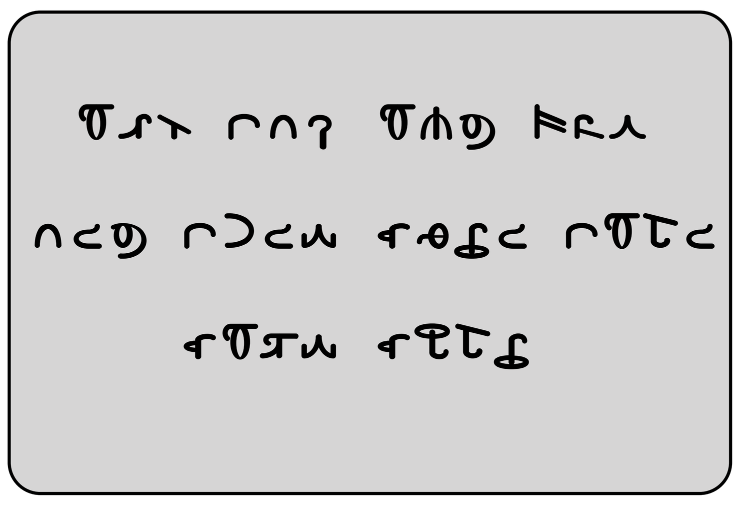
Experiment 2: "Prefixes" 69



PERCEPTIBILITY PARADIGM

1. MEMORISATION PHASE

- Duration: 3 minutes

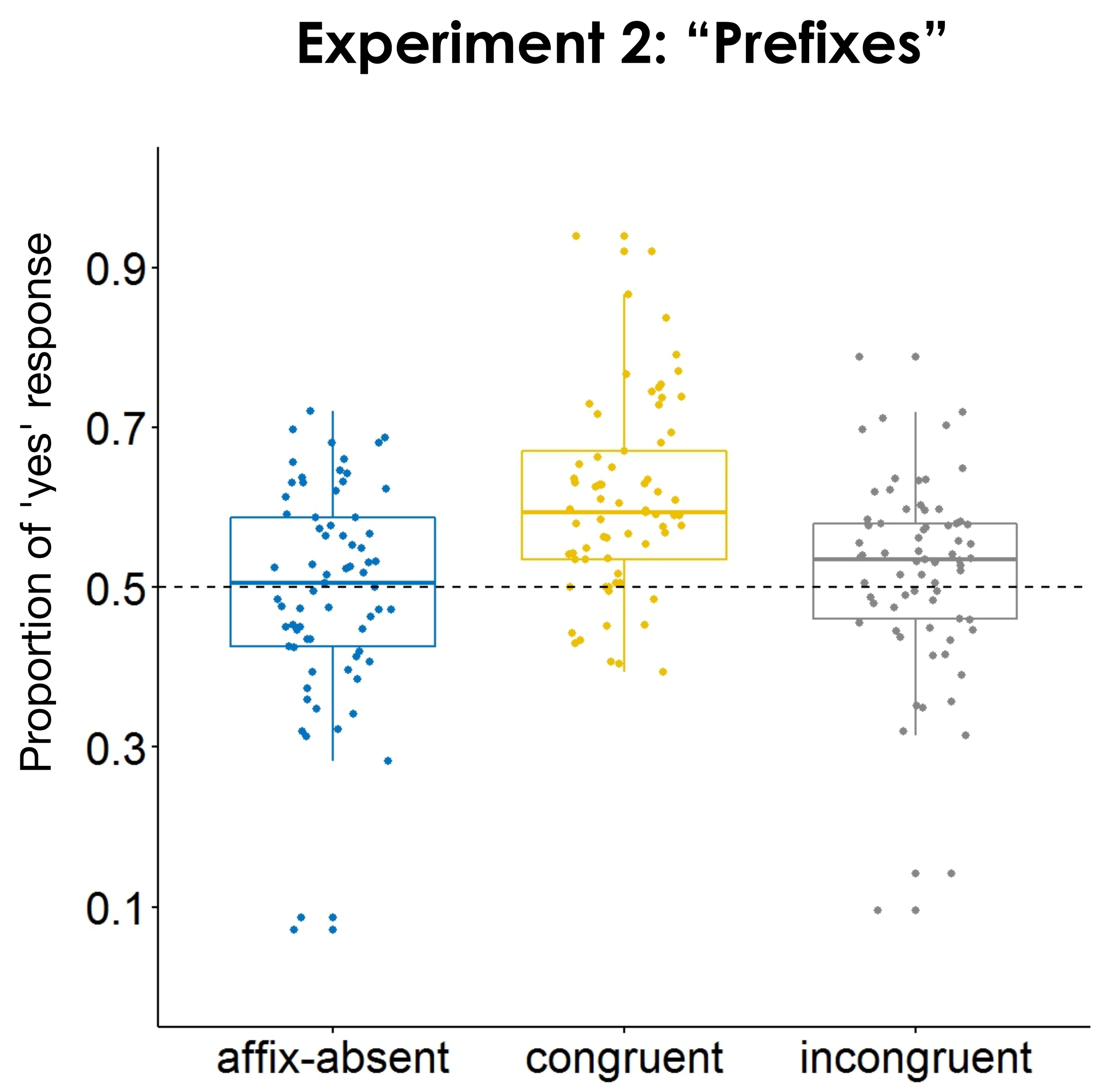


2. DETECTION TASK

- Novel pseudo-letter strings
- 40 strings with affix in initial position
- 40 strings with affix in final position
- 40 affix-absent strings

Note. Affix-like chunks are highlighted for illustration.

RESULTS



- Higher probability of 'yes' response for affix-present compared to affix-absent strings (Exp.1: $B(0.05) = 0.22$, $z = 4.02$, $p < .001$; Exp.2: $B(0.08) = 0.30$, $z = 3.64$, $p < .001$)
- Higher probability of 'yes' response for position-congruent compared to position-incongruent strings (Exp.1: $B(0.03) = 0.05$, $z = 1.86$, $p = .063$; Exp.2: $B(0.11) = 0.41$, $z = 3.85$, $p < .001$)
- Detection task: better detection for affixes in the initial position, indicating a string-initial perceptual advantage(Exp.1: $t(62) = 3.07$, $p = .003$, Cohen's $d = 0.39$; Exp.2: $t(67) = 3.94$, $p < .001$, Cohen's $d = 0.48$)

CONCLUSIONS

- When exposed to a large set of word-like items, readers spontaneously form representations for chunks of co-occurring characters and code for their typical position within these strings. Crucially, in the absence of any linguistic information, chunking relied only on the probabilistic information determining the internal structure of the novel words.
- Current findings provide evidence that morpheme-chunking during visual word processing can be, at least partly, ascribed to a general cognitive chunking mechanism that captures statistical regularities in the co-occurrence of visual objects [3,5].

REFERENCES:
[1] Amenta, S. & Crepaldi, D. (2012). *Frontiers in Psychology, 3, 232.* https://doi.org/10.3389/fpsyg.2012.00232; **[2]** Crepaldi, D., Rastle, K., & Davis, C. J. (2010). *Memory & Cognition*, *38*, 312–21.https://doi.org/10.3758/MC.38.3.312; **[3]** Fiser, J., & Aslin, R. N. (2005). *Journal of Experimental Psychology: General*, *134*, 521-537. http://dx.doi.org/10.1037/0096-3445.134.4.521; **[4]** Vidal, C., Content, A., & Chetail, F. (2017). *Behavior Research Methods, 49*, 2093-2112. https://doi.org/10.3758/s13428-016-0844-8; **[5]** Siegelman, N., Bogaerts, L., & Frost, R. (2017). *Behavior research methods*, *49*, 418-432. https://doi.org/10.3758/s13428-016-0719-z

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant Agreement No 679010 STATLEARN ERC-2015-STG/ERC-2015-STG.