

The role of prediction error in linguistic generalization and item-based learning

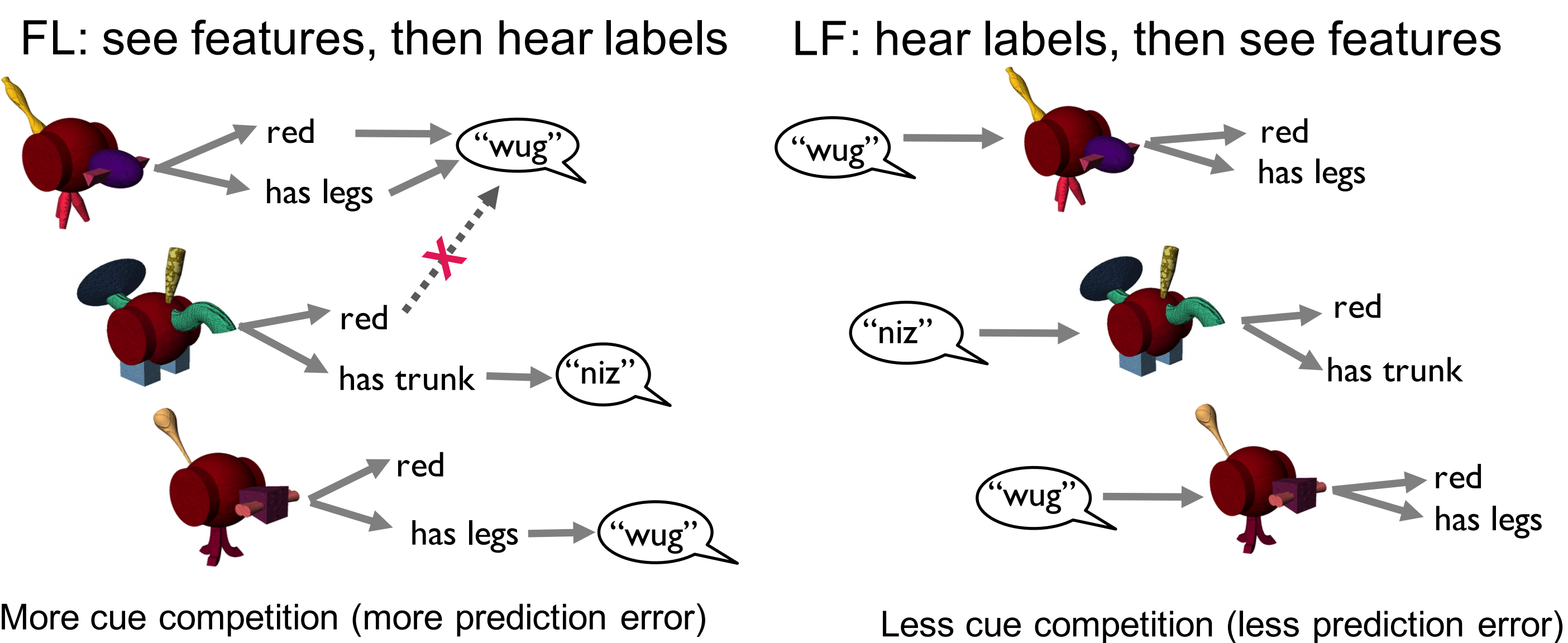
Maša Vujović¹, Michael Ramscar², & Elizabeth Wonnacott¹

¹ University College London, ² University of Tübingen



Prediction error drives (word) learning

Ramscar et al. (2010) found better word learning when there was more prediction error (FL learning):



Current study: prediction error & morphosyntax

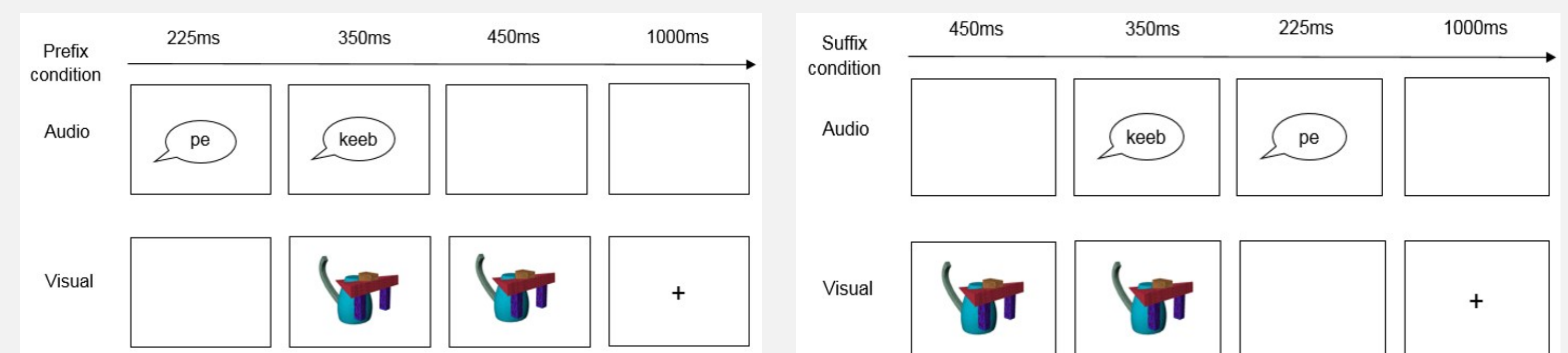
We teach participants an artificial language with affixation
Prefix condition: **ge meeb**
Suffix condition: **meeb ge**

	Category 1: ge			Category 2: ma		
75%						
25%						

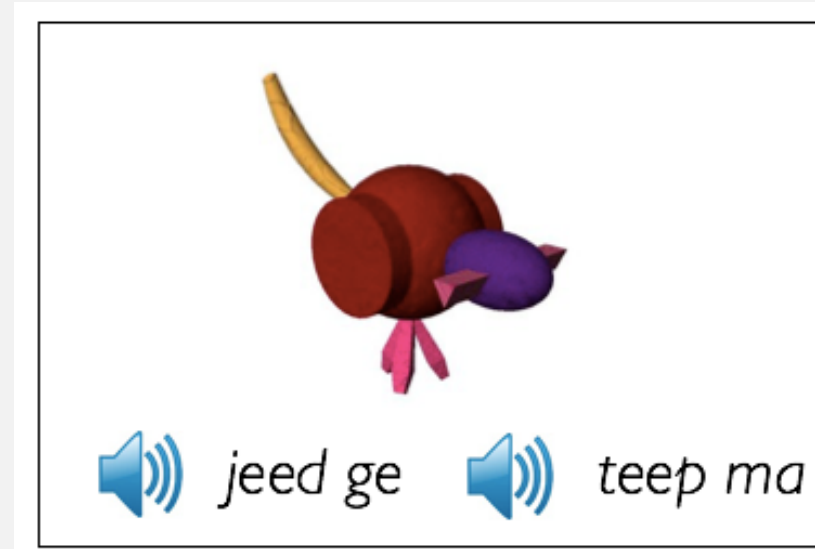
Prefix: better vocabulary learning
Suffix: better generalization of low type-frequency items

Experiment 1

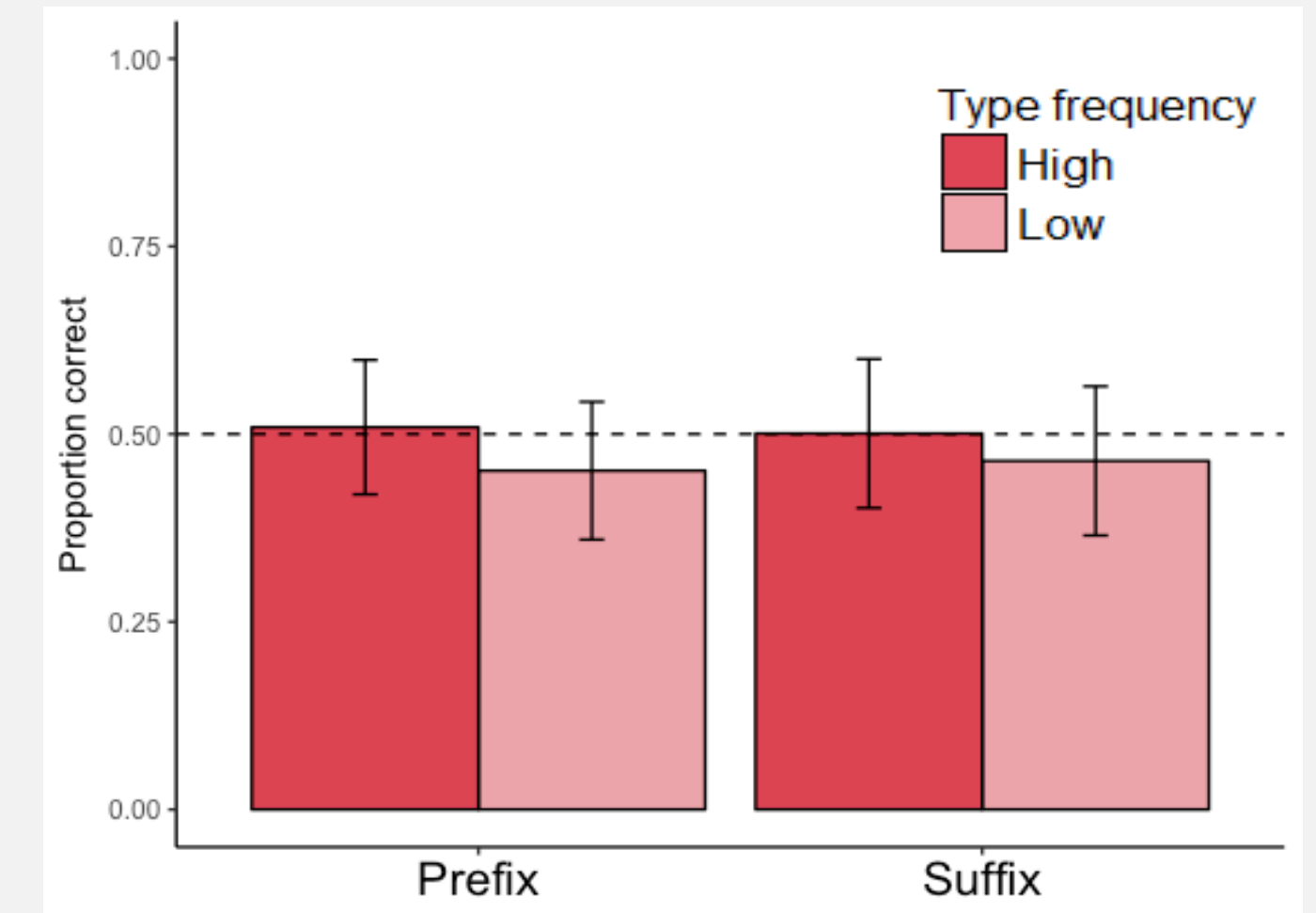
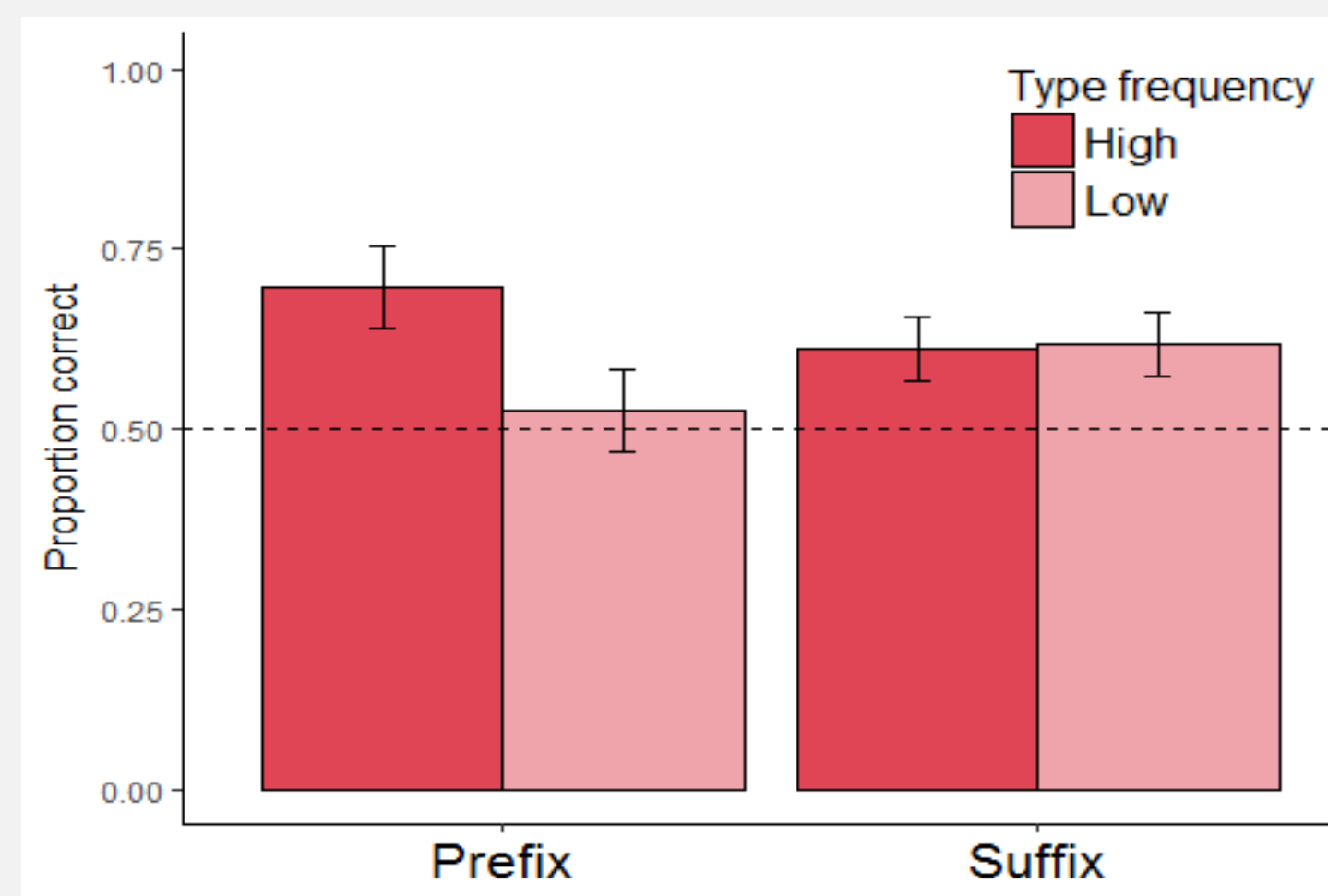
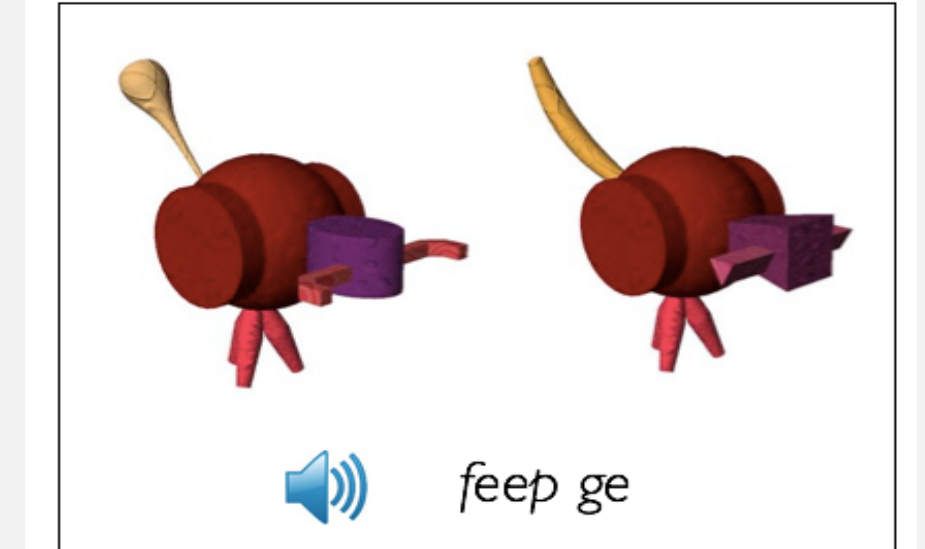
144 adult native English speakers recruited through Prolific Academic. **Training:** 8 items per category (6 HF, 2 LF), 16 exposures per item



Generalization



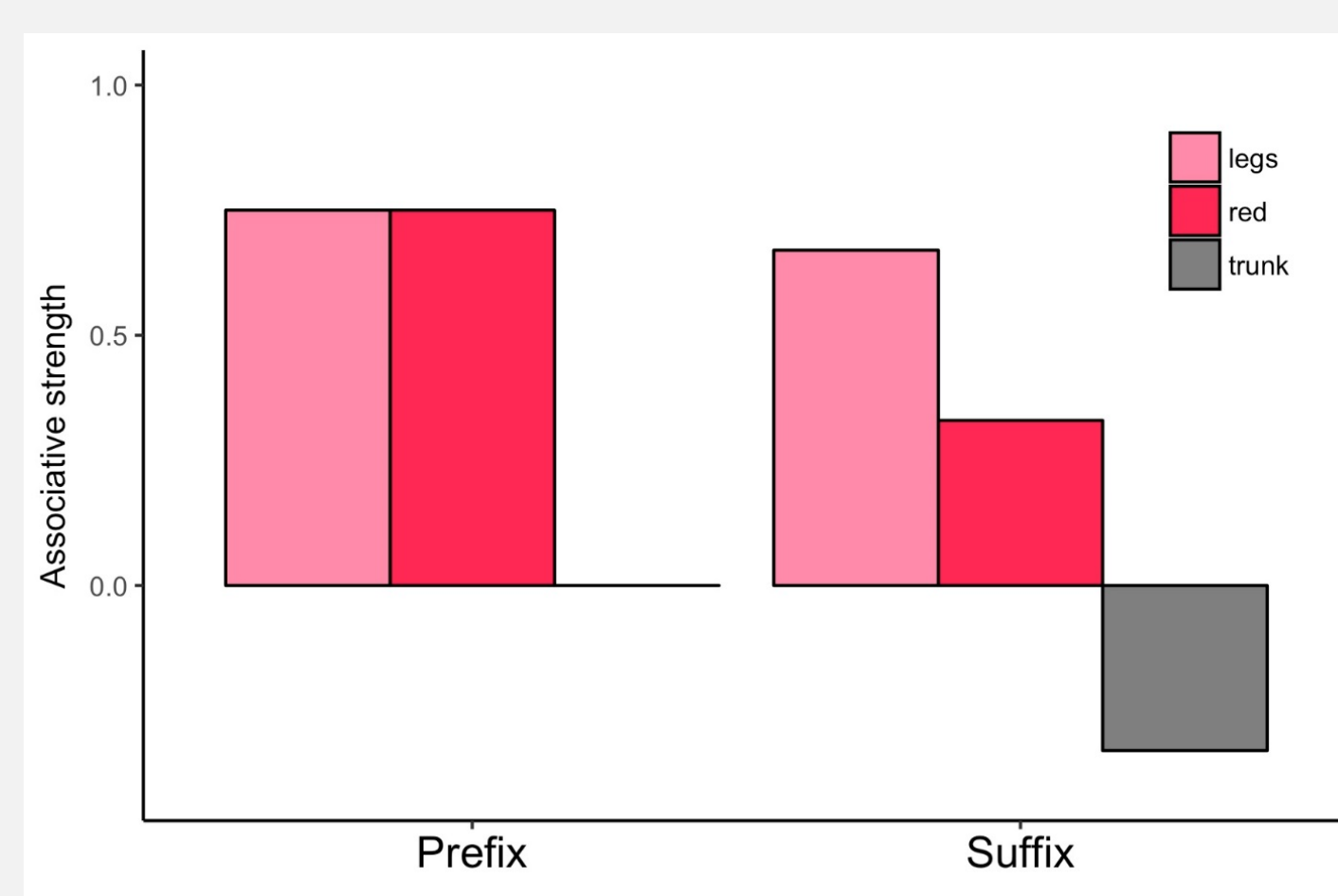
Vocabulary



Generalization: significant interaction between affix & type frequency. **Vocabulary:** no evidence of learning. Test too hard? Too many items?

Simulation experiment

Two neural networks trained on the artificial language with discriminative implementation of the delta rule (Widrow & Hoff, 1960).



Network	Probability of choosing correct category	
	HF	LF
Prefix	.99	0.5
Suffix	.99	.99

The maths:

$$(1) \Delta V_{ij}^n = \alpha_i \beta_j (\lambda_j - V_{TOTAL}) \quad \text{learning}$$

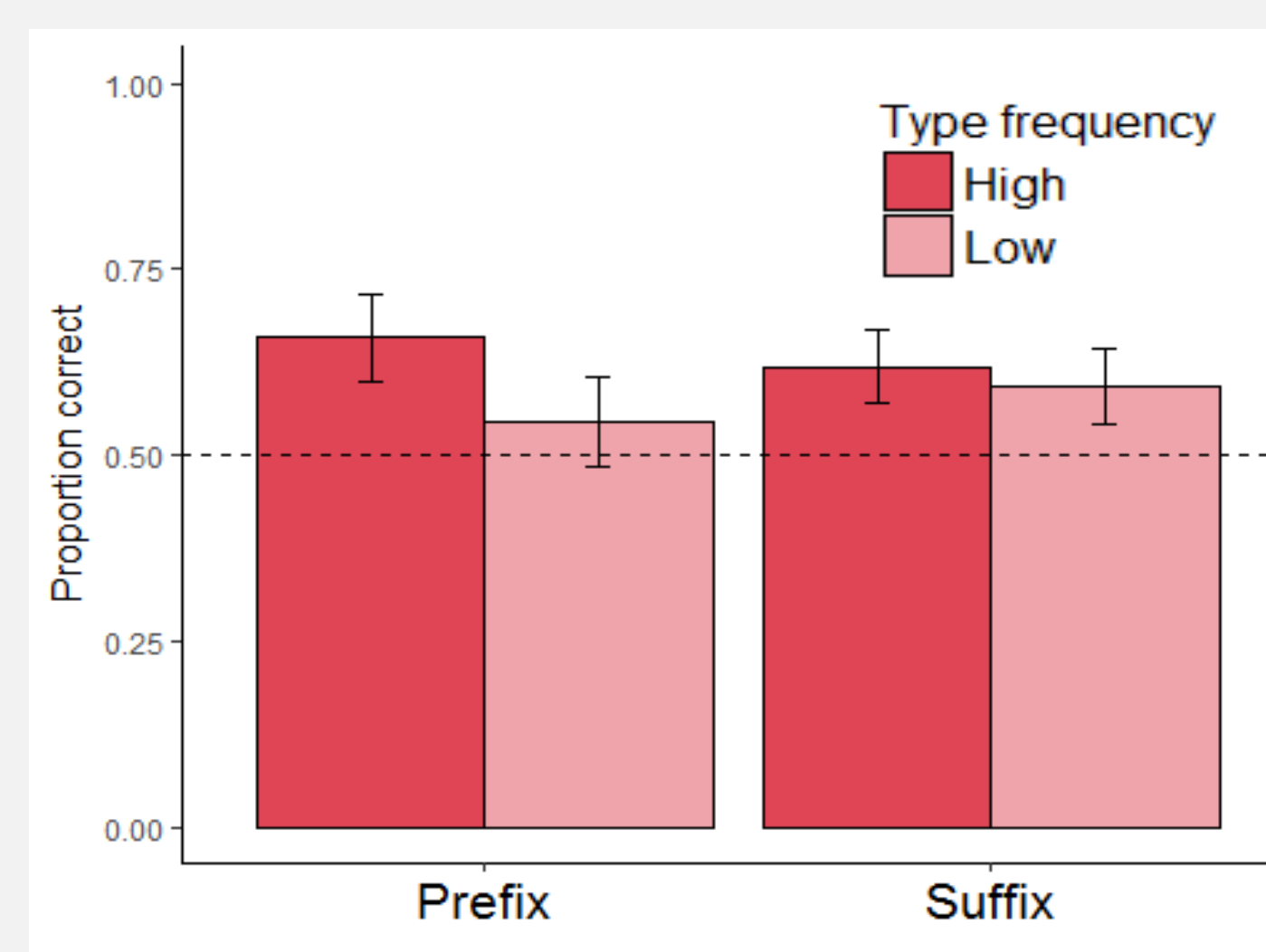
$$(2) V_{ij}^{n+1} = V_{ij}^n + \Delta V_{ij}^n \quad \text{update}$$

V – strength of association between set of cues i and outcome j
 α and β – learning rates; λ – maximum amount of associative value an outcome j can support; V_{total} – predicted response

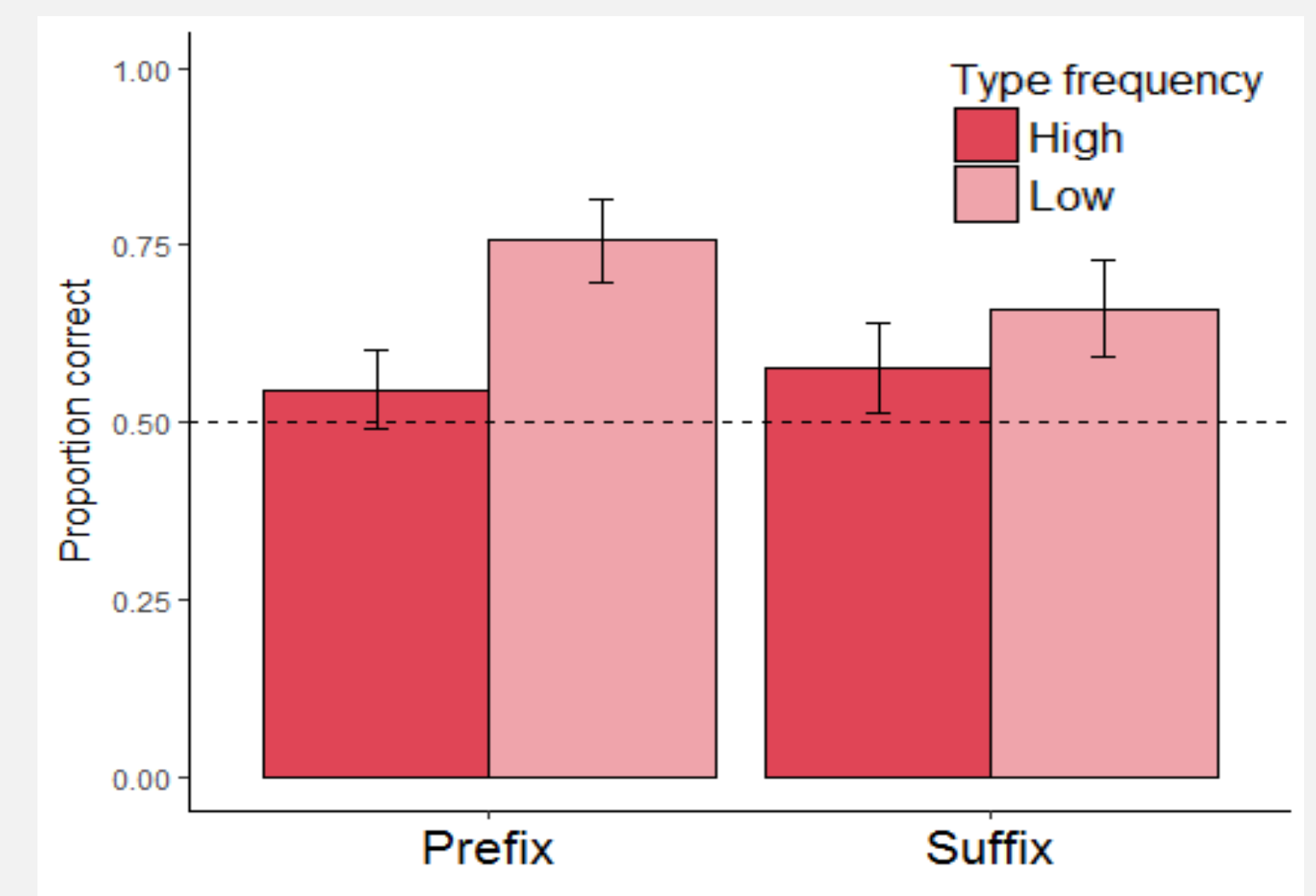
Experiment 2

170 adult native English speakers recruited through Prolific Academic. **Training:** 4 items per category (3 HF, 1 LF), 32 exposures per item. **Testing:** in vocab, foil is a novel item

Generalization



Vocabulary



Generalization: significant interaction between affix and type frequency (replicated Experiment 1). **Vocabulary:** no main effect of affix, but significant main effect of type frequency.

Summary & discussion

Generalization: interaction between affix and type frequency – only suffix generalize LF and HF items correctly, *as predicted*

❖ Support for the discrimination learning framework: the suffix condition learned predictive values of visual features for the affixes, and the prefix condition learned conditional probabilities of visual features given the affix.

Vocabulary learning: no prefix advantage, *contrary to prediction*

❖ Prefixing should make individual items more predictable (Ramscar 2013; Dye et al., 2016) – might need a different training procedure or more power to see a prefix advantage. In our study, vocabulary more learnable for low type frequency items, but this does not seem to help with generalization, speaking against the idea that item-based learning necessarily precedes generalization.

