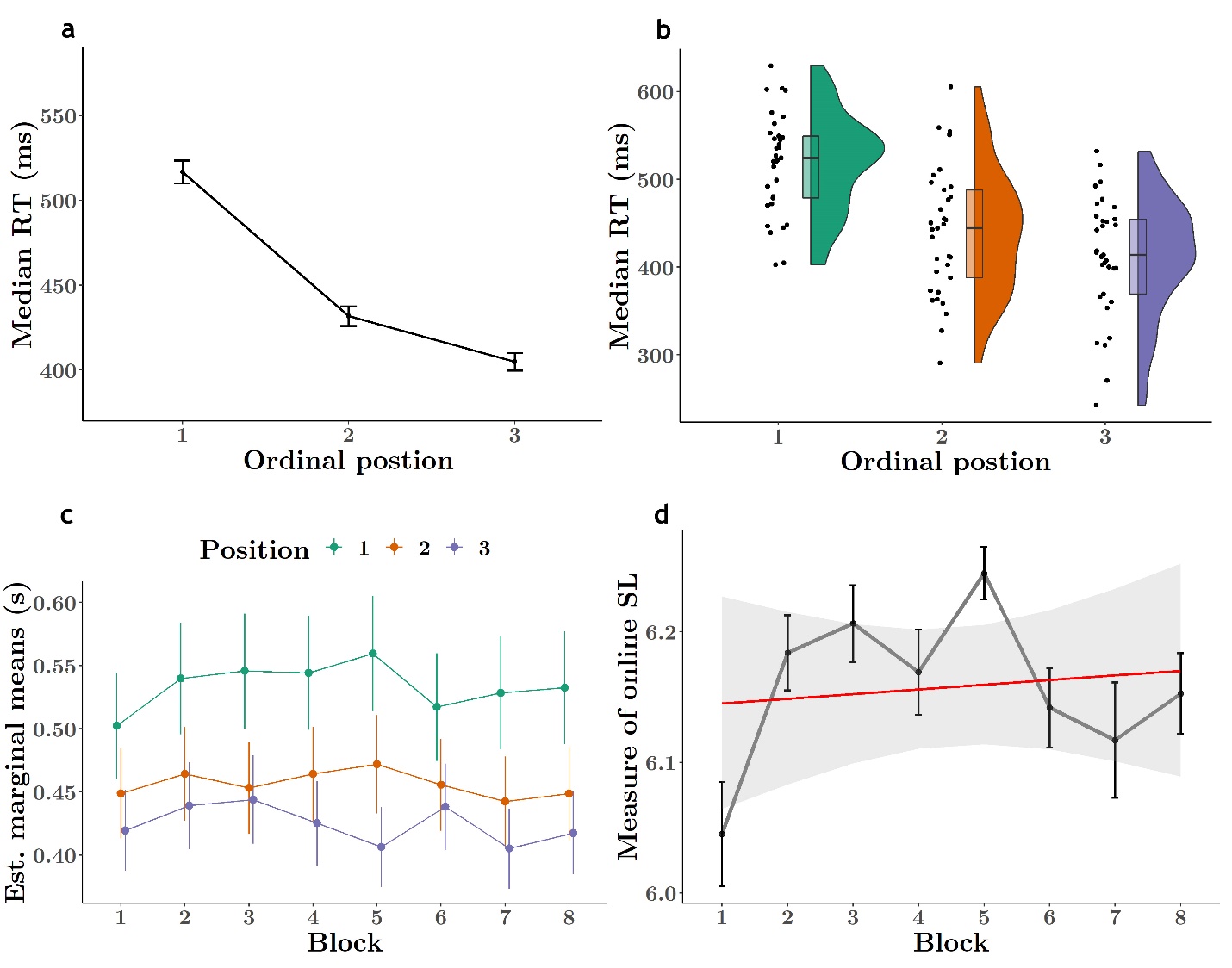
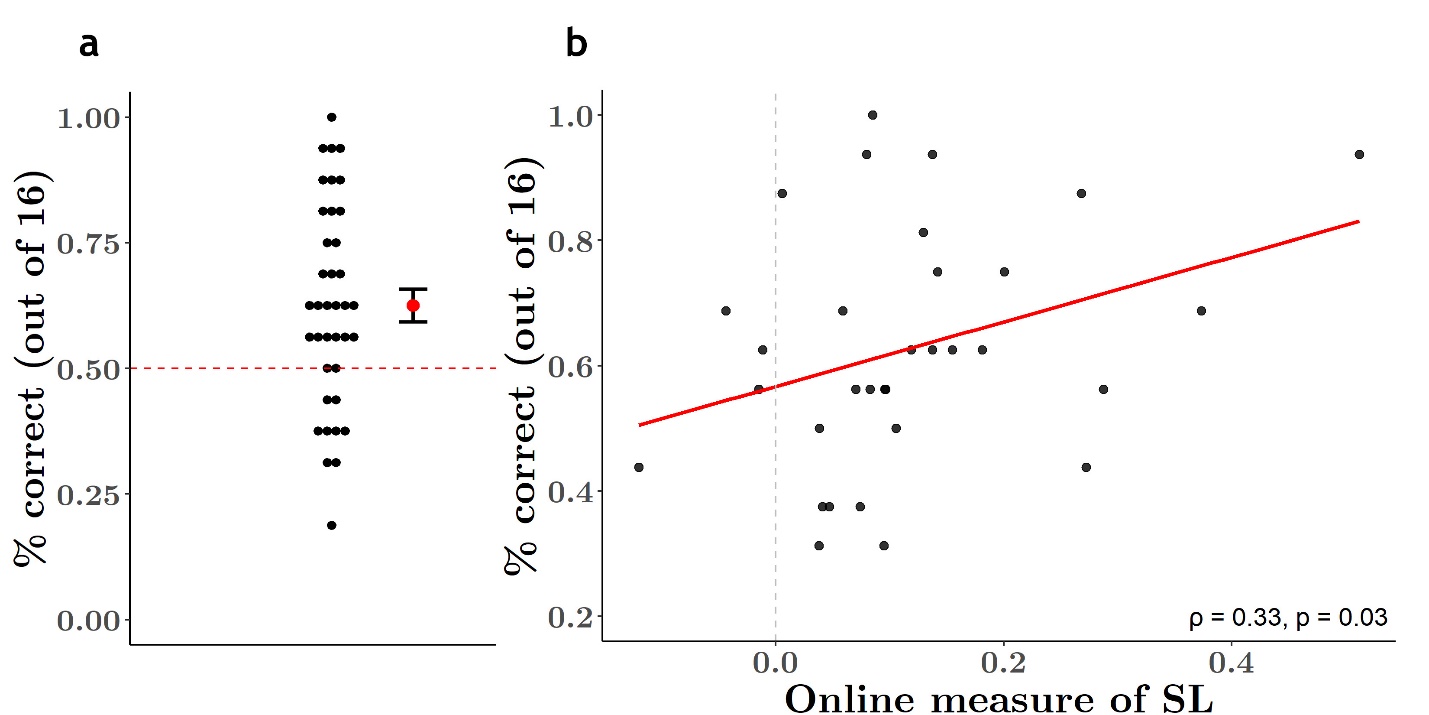
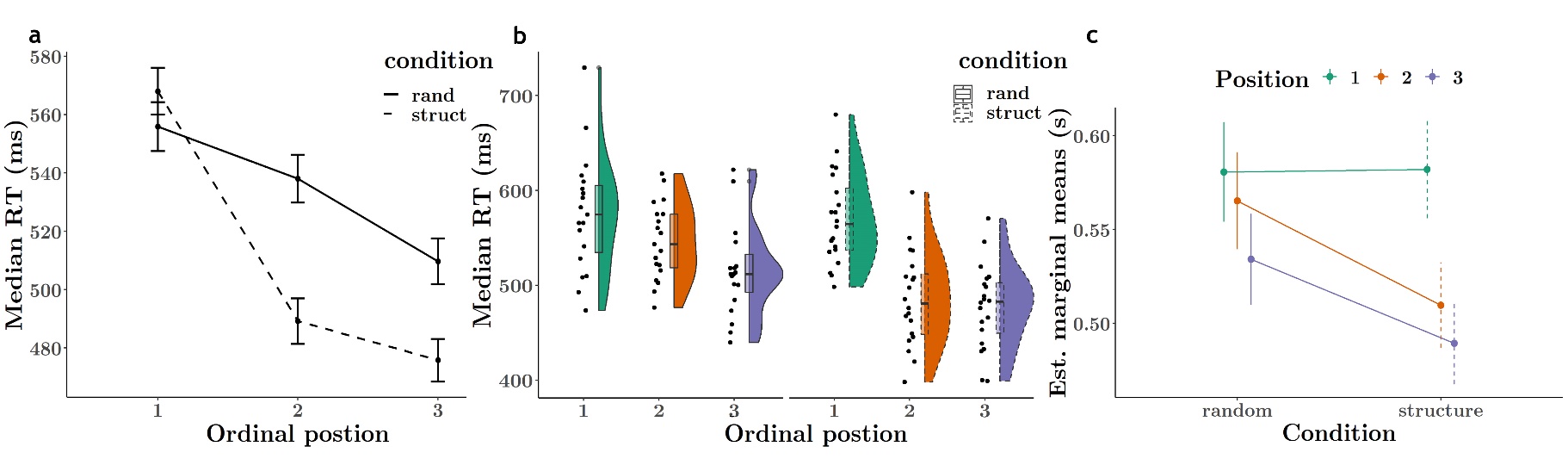
**Figures**



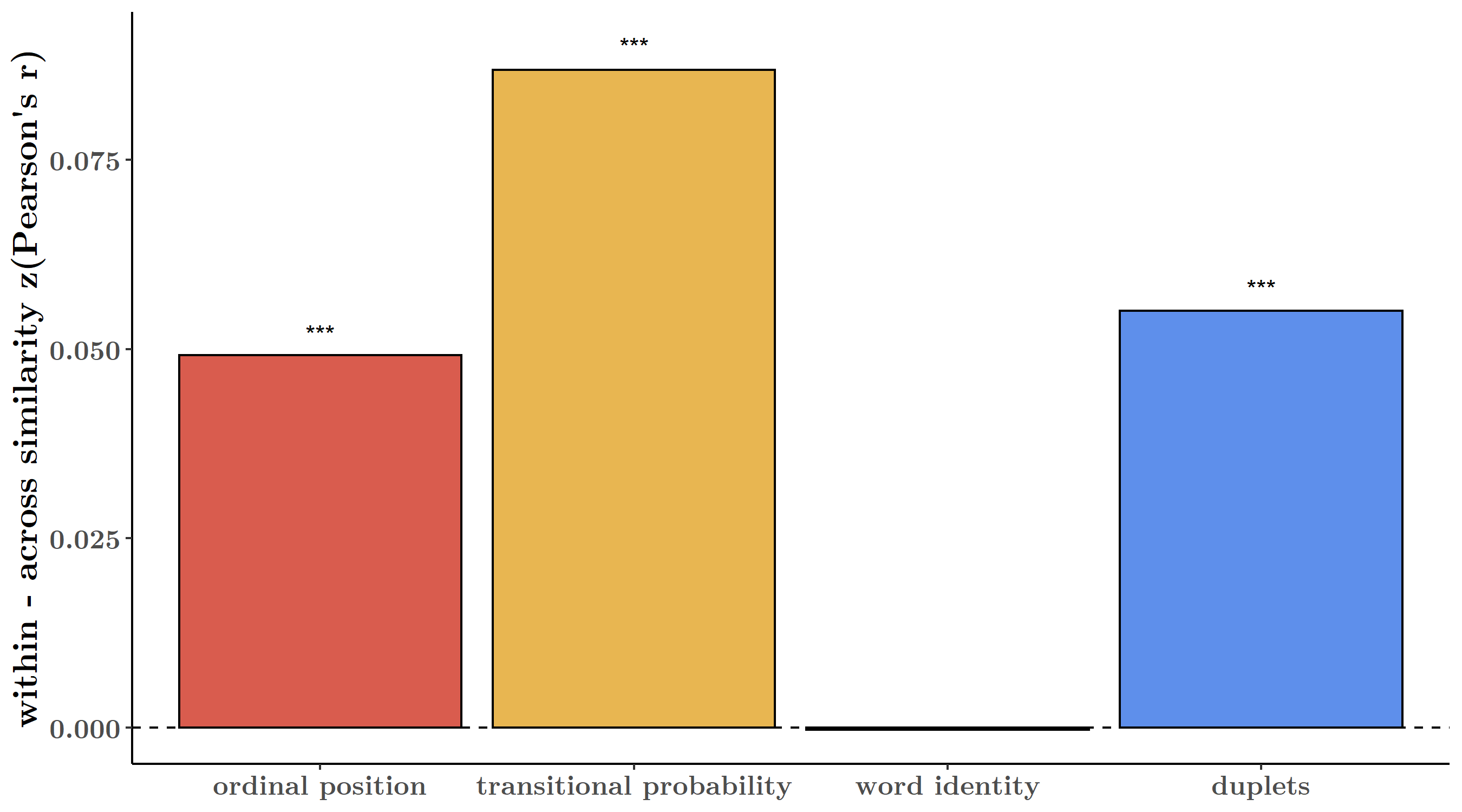
**Figure 1. Exp. 1: Online target detection reveals rapid and robust sensitivity to embedded regularities.** A. Median reaction times (RT) to target syllables are modulated by the syllable’s ordinal position in pseudowords. Participants responded more slowly to syllables in the word-initial (1st) position than to syllables in the word-medial (2nd) or word-final (3rd) position. Error bars represent 95% confidence intervals. B. Distribution of median RTs to each ordinal position for each participant (black dots). (Jittered along x-axis for visibility.) Box plots indicate group median and 95% CI. C. RTs to targets in the 1st vs 2nd or 3rd position are somewhat distinct in the first block (first 3 minutes of exposure), and clearly differentiated thereafter. Points represents estimated marginal means from the GLMM, vertical lines represent 95% CI. D. The magnitude of the SL effect (log mean RT to 1st position – log mean RT to (2nd & 3rd) position syllables) was smallest in the first block, but rose dramatically in the second block, thereafter hovering around the overall mean. (No difference in mean RTs to the three positions would result in a value of 0.) We observed no significant effect of block on this metric, suggesting that the modulation of RT as a function of target syllable position in the pseudowords had already occurred in the first block.



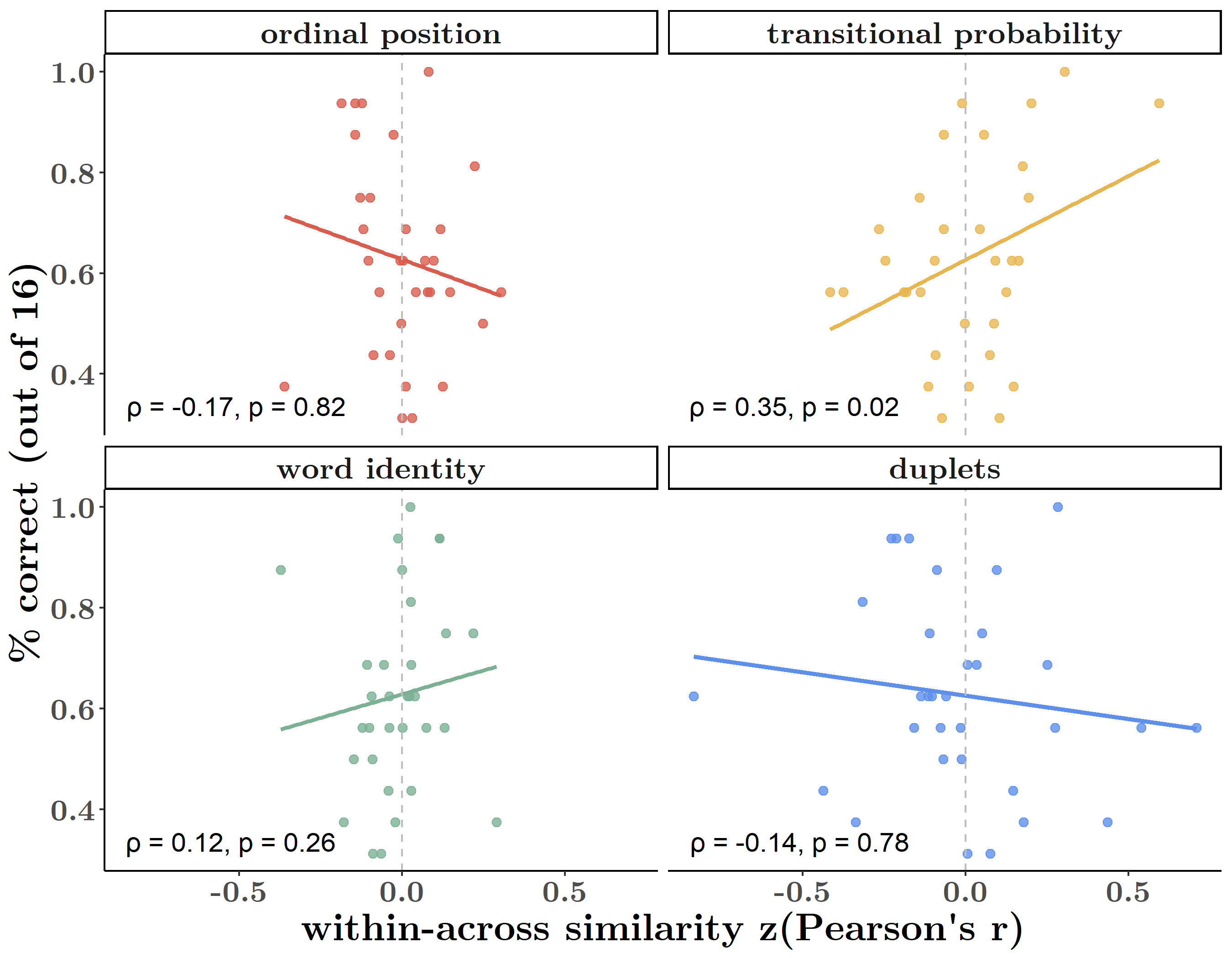
**Figure 2. Exp. 1: Participants discriminated pseudowords from part-words, but online and offline SL was only weakly correlated.** A. Word recognition performance was above the 0.5 chance level (dashed red line), suggesting participants were able to use implicitly learned regularities during the exposure phase to explicitly discriminate pseudowords from part-word foils. Red dot and error bars represent group mean (0.62) and SEM (0.2). Black dots represent percent correct trials (out of 16) for each individual. B. Pearson correlation of each individual’s word recognition performance (percent correct trials) and RT effect in the target detection task (online measure of SL). Performance in the two tasks was weakly correlated (). Dashed grey line represents threshold at which there was no difference in mean RT to targets in the three positions: those with values above 0 showed faster responses to 2nd and 3rd position targets vs. 1st position targets, those with values below 0 showed the opposite effect.



**Figure 4. Exp. 2: RT to target syllables modulated by ordinal position and the presence of structure.** A. In the structured condition (dashed line), the ordinal position of target syllables modulated RT, replicating Exp. 1. Surprisingly, a less pronounced RT modulation by position was also observed in the random condition (solid line). Error bars represent 95% confidence intervals. B. Distribution of median RTs to each ordinal position for each participant (black dots) in the random condition (left, solid violin & boxplot outline) and structured condition (right, dashed violin & boxplot plot outline). (Jittered along x-axis for visibility.) Box plots indicate group median and 95% CI. C. Estimated marginal means from GLMM with ordinal position and condition as predictors for RT. RTs to target syllables in the 1st position are equal between conditions, but RTs to those in 2nd and 3rd positions are significantly faster in the structured as compared with the random condition. This suggests the transitional probability structure was responsible for the RT effect, since less predictable syllables (1st position, TP = 0.33) were responded to equally quickly, but predictable syllables (2nd and 3rd position, TP = 1), which appeared only in the structured stream, elicited markedly faster responses.

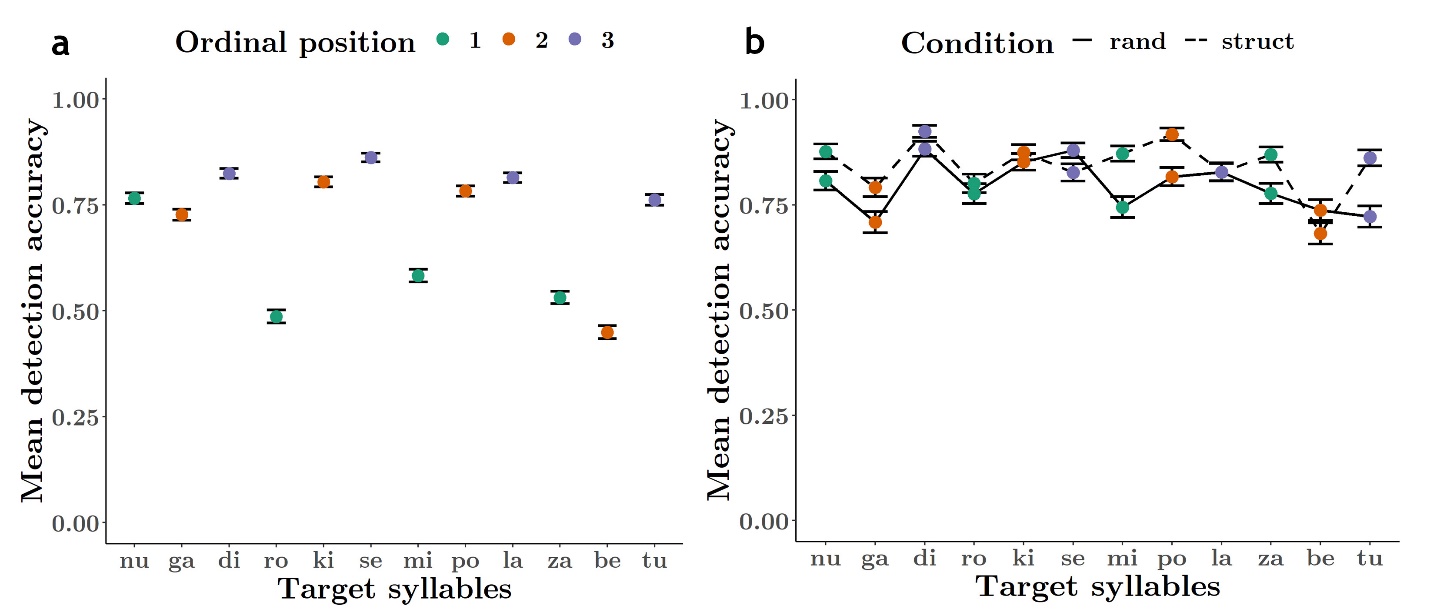
****

**Figure 5. RTs reveal representational similarity for ordinal position, transitional probability, and duplet pairings, but not pseudoword chunks.** Representation similarity analysis showed that groups of RTs classified as belonging to the same category of ordinal position, transitional probability, and duplet pairings (within a given pseudoword) were more similar than groups of RTs that spanned different categories. Within similarity minus across similarity was greater than zero for three of four features that describe the structured speech stream. (Wilcoxon’s rank sum test for paired groups, on bootstrapped z-transformed Pearson’s correlations between syllables for each participant.) \*\*\*p<0.0001

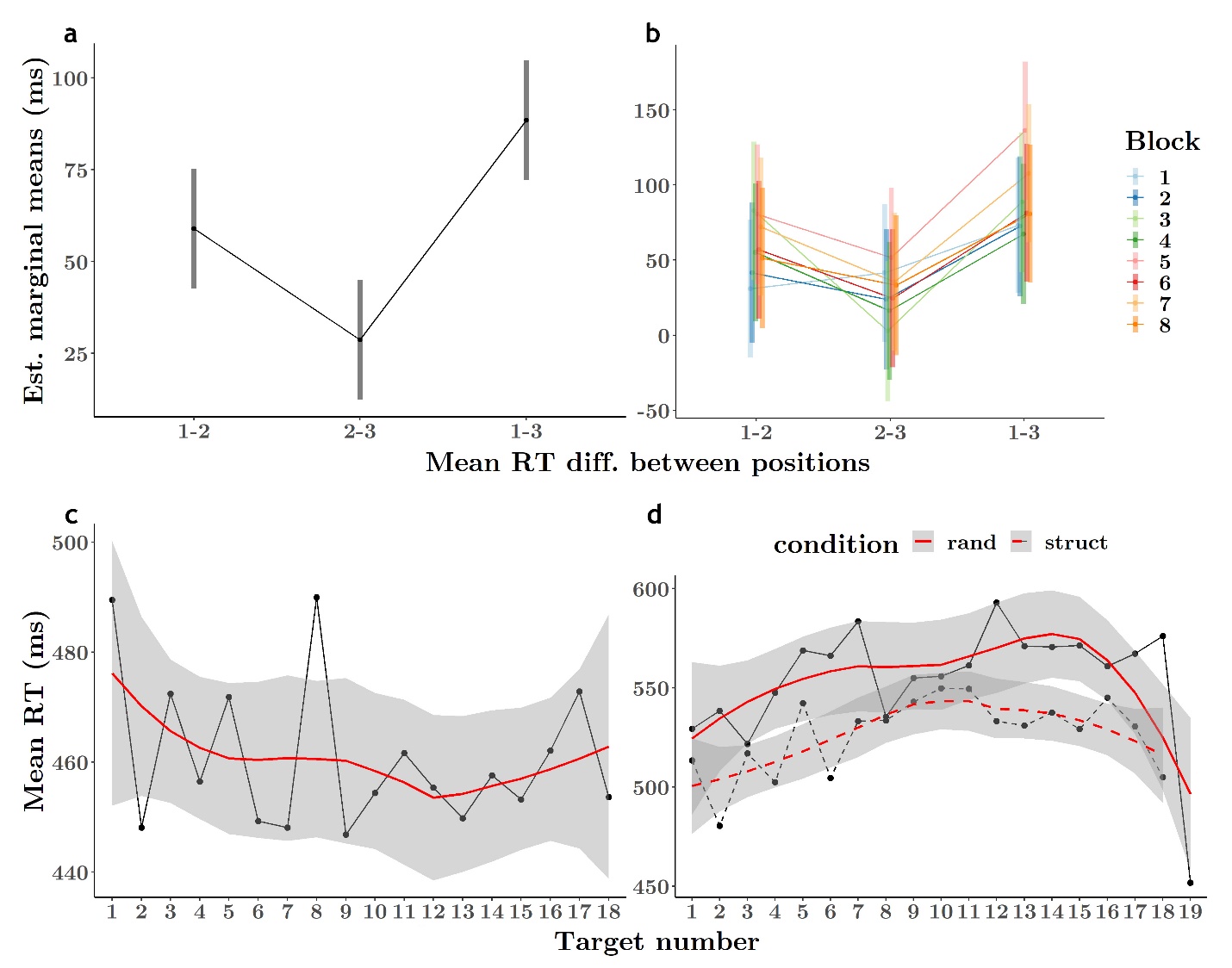
****

**Figure 6. Sensitivity to transitional probability, but not other features, correlates with word recognition performance.** Word recognition performance (% correct trials) correlates weakly with the tracking (measured as within-across similarity in RT) of ordinal position, duplet pairings, and pseudoword identity, however it reveals a small but significant correlation with the tracking of transitional probabilities. (One-sided t-test on Pearson’s product-moment coefficient.)

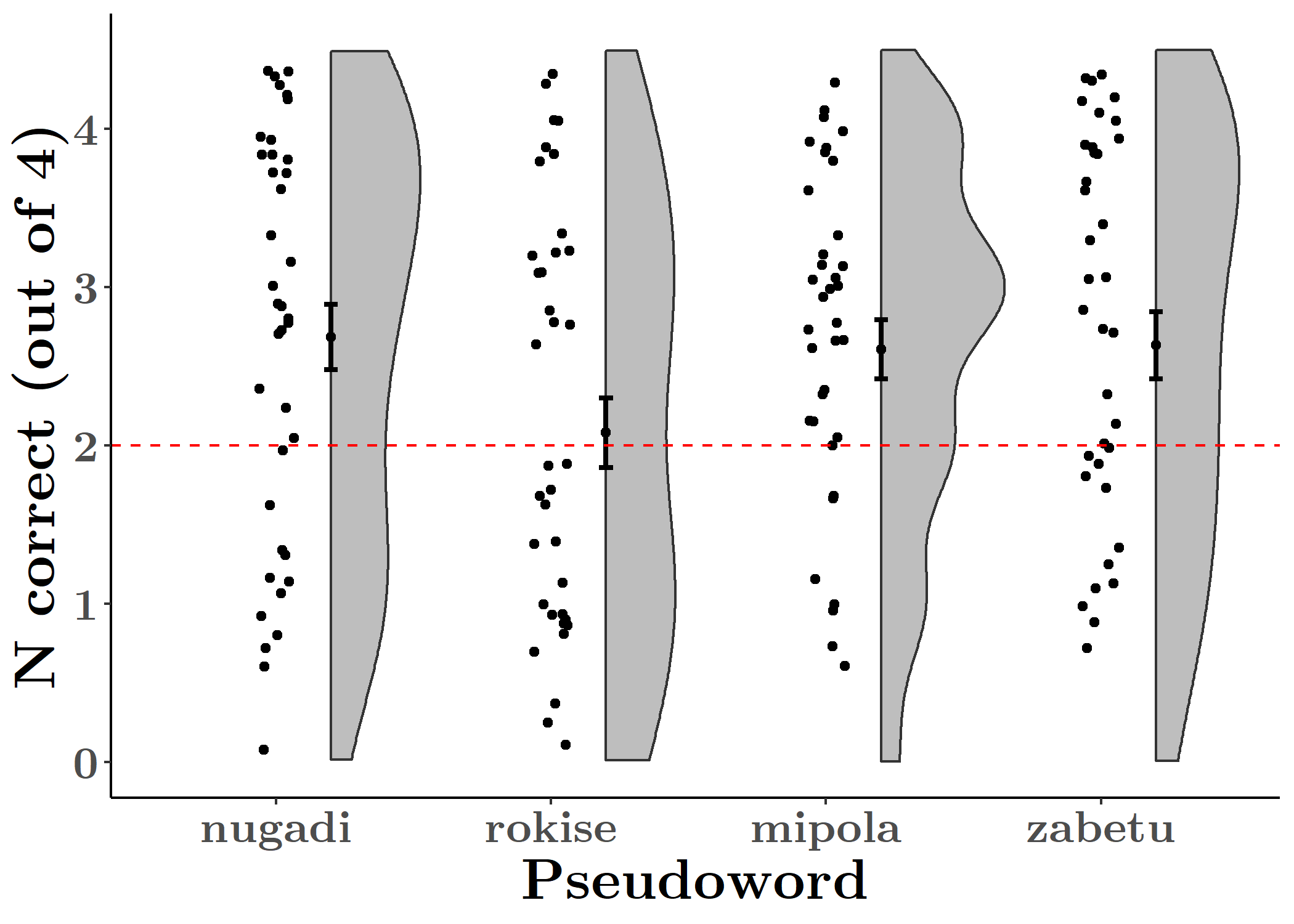
**Supplementary Figures**

****

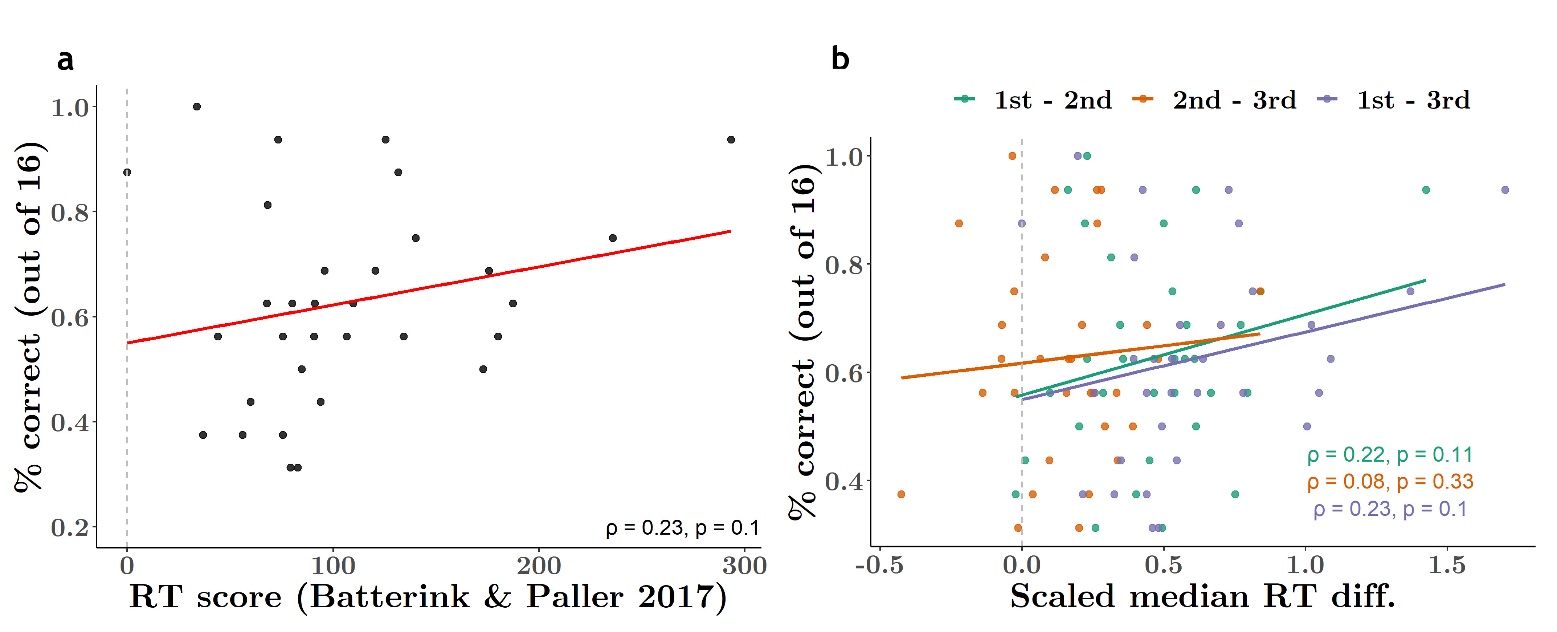
**Figure S1. Detection accuracy in Experiments 1 & 2.** A. Overall detection accuracy (hit rate) for each target syllable in Experiment 1. Colors indicate ordinal position of each syllable. Accuracy was generally lower for 1st position syllables vs. 2nd and 3rd position syllables. B. Overall detection accuracy (hit rate) for each target syllable, in each condition, in Experiment 2. Error bars represent SEM. Accuracy was higher for target syllables in the structured condition versus the random condition.



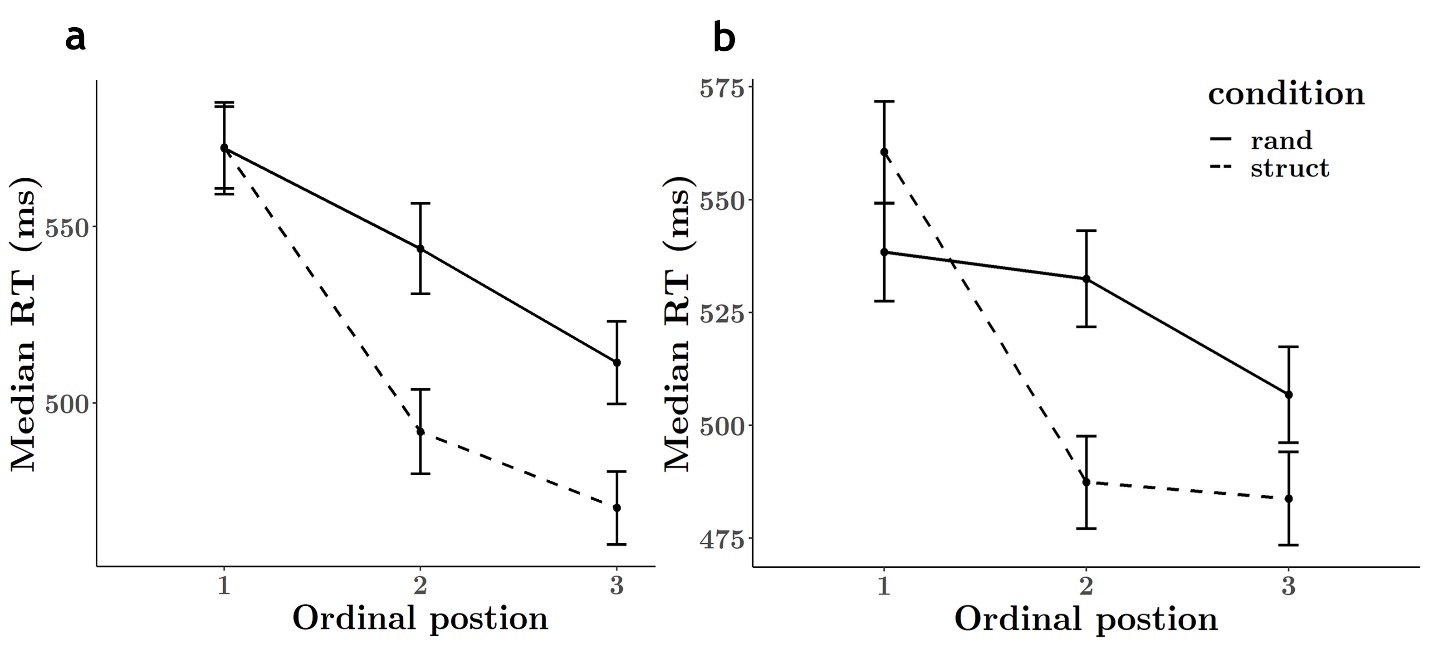
**Figure S2. Exp. 1: Controls to disconfirm effect of spurious RT facilitation.** A. The mean difference in RT between 1st and 2nd position targets was higher than the mean difference between 2nd and 3rd position syllables (the 1st-3rd RT difference is greatest). Graded RT effect is not linear, and therefore cannot be ascribed simply to faster RTs in later stream positions. Black dots are estimated marginal means from linear model with block and pairings as predictors. Vertical lines represent 95% CIs. B. Mean RT differences between pairs of positions did not vary significantly by block, suggesting the general RT effect did not vary much over longer exposures. Center dots are estimated marginal means from the linear model, vertical lines represent 95% CIs. C. Exp. 1: RTs to targets in each stream position (averaged over ordinal positions and blocks) did not show a consistent downward trend. Dots represent mean RTs to targets in each stream position, red line indicates linear prediction, grey ribbon indicates SEM. D. Exp. 2: RTs to targets in each stream position for each condition (averaged over ordinal positions and trials) showed a quadratically-shaped trend that increased in overall RT until target number 12-14, and thereafter began to decrease. This pattern is unlikely to have generated the graded RT effect we observed in our main RT by ordinal position analysis. Dots represent mean RTs to targets in each stream position, red line indicates linear prediction, grey ribbon indicates SEM.

****

**Figure S3. Exp. 1: Preference for pseudowords over part-word foils was above chance for 3 of 4 pseudowords.** Number correct trials (out of 4 for each pseudoword) was above chance (2) for three of the four pseudowords. Jittered dots represent number correct trials for each individual. Adjacent black dot and error bars represent group mean and SEM. Overall word recognition performance was therefore not driven by successful discrimination of merely word.



**Figure S4. Exp. 1: Correlation between online and offline SL measures following two exploratory methods reveals no significant relationship.** A. Word recognition (% correct) was correlated for each individual with an RT score, as per Batterink & Paller 2017, calculated by simple subtraction of median 1st-median 3rd position RT (Pearson’s product-moment correlation). We found a weak correlation (). B. Word recognition was correlated with three measures of online SL for each participant: we computed the difference in scaled median RTs between ordinal position separately. Scaling median RT values allowed all participants’ scores to be more comparable. Calculating median differences for each pair allowed us to account for the possibility that e.g. the difference between 1st and 2nd positions better predicted word recognition performance than the difference between 2nd and 3rd position RTs. Median RT differences between all three pairs had only a weak relationship with word recognition performance. The strongest relationship existed between word recognition accuracy and 1st-3rd median RT (), which revealed the exact same relationship as Batterink & Paller’s method. Dashed grey line in both graphs represents threshold at which there was no difference in RTs between the respective position pair.

****

**Figure S6. Exp. 2: Graded RT effect for random and structured conditions in both between-subject condition orders.** A. Overall median RTs to targets in each ordinal position for participants who performed the structured condition first and the random condition second. B. Overall median RTs for participants who performed the random condition first and the structured condition second. Error bars represent 95% CIs. RT is strongly modulated by ordinal position in the structured condition in both between-subject orders. Those who performed the random condition first showed somewhat smaller differences in their RTs to each position (B, solid line) than those who performed this task after exposure to the structured stream (A, solid line). Despite observing faster responses to later ordinal positions in the random condition, the overall RT effect can be safely ascribed to the embedded regularities, as the RT effect is significantly larger in the structured condition (A-B, dashed line).