DO ERP INDICES OF RATIONAL INFERENCE IN REAL-TIME COMPREHENSION TRACK THE NOISE RATE IN THE ENVIRONMENT?

Jiaxuan Li (UC Irvine), Alyssa Ortega (UC Merced), Richard Futrell (UC Irvine), Rachel Ryskin (UC Merced)
jiaxuan.li@uci.edu

Introduction According to the Noisy-Channel framework [1,2], the communication process involves noise (e.g., speaker errors, mishearing, noisy environment) and comprehenders rationally infer the intended meaning given imperfect input. In offline tasks, inference about non-literal interpretations is sensitive to statistical patterns in the linguistic environment [1, 2, 3]. Comprehenders are more likely to arrive at a non-literal interpretation when there is a possible intended message that is likely to be distorted into the received message; and (ii) the perceived noise rate of the communicative environment is higher, and this inference adapts rapidly to the rate and characteristics of noise in the environment [3]. In online real-time processing, the N400 and P600 ERP components seem to index the probability of noisy-channel inference. When an implausible continuation is orthographically close to a more plausible alternative, an increased P600 and decreased N400 are observed [4, 5]. If this view of P600 is correct, then it should also adapt rapidly to the noise rate of the linguistic environment in online rational inference. The current study replicates and extends [4] to investigate whether N400 and P600 effect sizes are modulated by the noise rate by manipulating the proportion of errors in non-critical exposure sentences. Data collection is ongoing (planned N = 48).

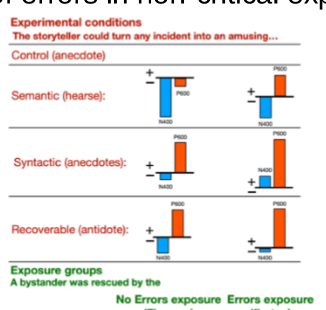

Fig.1. Illustration of the experimental design and predictions according to the Noisy-Channel framework

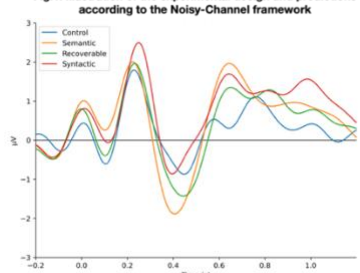

Fig.2. Grand-averaged ERP waveforms to critical nouns across centro-posterior electrodes.

Experimental Design We crossed error type conditions (within-subjects) with exposure groups (between-subjects) (Fig. 1). **Error type:** In the *semantic* condition, the last word is semantically implausible (e.g. The storyteller could turn any incident into an amusing hearse); in the *syntactic* condition it contains a *syntactic* error (e.g. anecdotes). In the *recoverable* condition, it is semantically anomalous but orthographically close (as measured by Levenshtein distance) to a semantically plausible neighbor (e.g. antidote). The errors in *syntactic* and *recoverable* conditions are both recoverable via a noisy-channel inference and expected to elicit a P600. **Exposure:** We also manipulated the rate of errors that comprehenders are exposed to. In the *Errors* exposure group, exposure sentences contain blatant typographical errors (e.g. A bystander was rescued by the firetan). In the *No Errors* exposure group, exposure sentences did not contain errors.

Materials We adapted stimuli from [4]. There are 640 experimental sentences with 160 in each condition. There are 40 exposure sentences for each group, in addition to 280 plausible filler sentences. All sentences are distributed into 8 lists following a Latin Square design, and the orders are randomized. Comprehenders are asked to read sentences presented word by word while EEG signals are recorded.

Analysis We analyze mean N400 (300-500ms post-onset) and P600 (600-800ms) amplitudes elicited by critical words over fifteen centro-parietal electrodes. We analyze the results in a linear mixed effect model: *Amplitude ~ Condition * Exposure + (1+Condition | subject) + (1+Condition*Exposure | item) + (1+Condition*Exposure | electrode)*

Predictions and Preliminary results We expect to replicate [4], where semantic and recoverable conditions elicited N400 effects, and syntactic and recoverable conditions elicited P600 effects (see Fig1-left). Critically, the Noisy-Channel framework predicts an interaction between exposure and conditions: The N400 effect in critical conditions (*semantic*, *recoverable*) will be greater when the noise rate is lower (*No Errors* exposure group), whereas P600 effects (in *syntactic* and *recoverable* conditions) will be greater when the noise rate is higher (*Errors* exposure group) (Fig1-right). In our preliminary analysis (N=16), we collapsed data from both exposure groups and replicated a significant N400 effect to semantic violations ($t = -2.03^*$) and a near significant P600 effect to syntactic violations ($t = 2.11$, $p = 0.06$). Comparison between exposure groups is forthcoming.

References: [1] Gibson et al. (2013) PNAS. [2] Levy, 2008. EMNLP. [3] Ryskin et al. (2018) Cognition. [4] Ryskin, Stearns et al. (2021). Neuropsychologia. [5] Li & Ettinger (2023). Cognition.