Resolving different types of prediction violations during sentence comprehension: Insights from pupillometry and individual differences

Kayla Keyue Chen, Chen Lu, Xinyi Wang, Yuxuan Wu, Wing-Yee Chow

University College London

Much evidence has shown that contextual information can facilitate retrieval and integration of a predicted word during sentence comprehension [1-2]. However, resolving prediction violations – whether unexpected-yet-plausible, which may be integrated through reinterpretation, or implausible, which introduce irreconcilable conflicts – may engage distinct cognitive processes [3-5]. This study investigated the resolution of different types of prediction violations using pupillometry and found that different executive functions predict individuals' pupillary responses to plausible and implausible violations.

Participants (n=57) listened to Mandarin Chinese sentences such as "*Xiao Mei is practicing dancing in the activity room, and in front of her, is a …*" continued with a critical word that was either expected (e.g., *mirror*), unexpected but plausible (e.g., *water flask*), or implausible (e.g., *forest*) while their pupil sizes were recorded. They rated the plausibility of each sentence on a scale from 1 (very implausible) to 5 (very plausible). Subsequently, we asked participants to complete 5 behavioural tasks to evaluate individual differences in conflict monitoring and cognitive control (Stroop, AX-CPT), working memory (backward-digit span, phonological span), and verbal fluency (semantic category).

We analysed changes in participants' pupil size from the onset of the critical noun to 2500 ms following noun onset. Pupil dilation was observed to be the most pronounced for implausible words, followed by unexpected words, while expected words elicited pupil contraction (Fig. 2). We applied growth curve analysis with a cubic orthogonal polynomial and extracted individual effect sizes. Backward stepwise multiple regressions revealed that the AY-AX cost significantly predicted the effect sizes of the Unexpected vs. Expected contrast on the intercept and quadratic term, suggesting that better conflict monitoring and cognitive control ability enhances sensitivity to plausible prediction violations (Fig. 3C). For the Implausible vs. Unexpected contrast, phonological span and verbal fluency were associated with the effect sizes on the intercept and cubic terms respectively. These associations suggest that individuals with higher verbal fluency are more sensitive to semantic anomalies, while those with lower phonological working memory find such violations more challenging to resolve (Fig. 3A & 3B).

These findings indicate that comprehenders may engage distinct cognitive mechanisms to process prediction violations that are plausible vs. implausible. Resolving plausible violations may require rapid conflict detection and resolution, whereas resolving implausible violations may rely more on rapid detection of semantic anomalies and effective maintenance of sentence context in working memory.
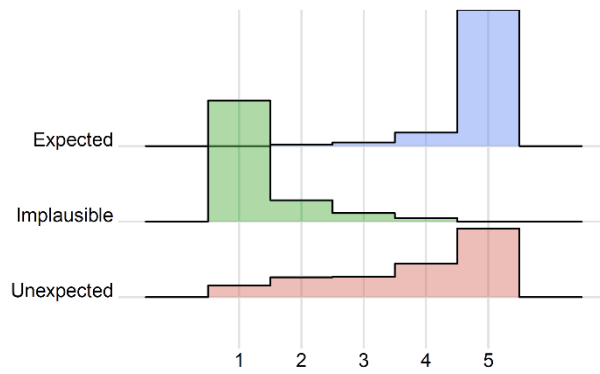
Figure 1. The distributions of plausibility rating (1-very implausible, 5-very plausible) in three conditions.

Reference
[1] Kamide, Y. (2008). *Language and Linguistics Compass*, 2(4), 647–670.
[2] Kutas, M., DeLong, K. A., & Smith, N. J. (2011). In *Predictions in the brain: Using our past to generate a future*. Oxford University Press.
[3] Van Petten, C., & Luka, B. J. (2012). *International Journal of Psychophysiology*, 83(2), 176–190.
[4] Ness, T., & Meltzer-Asscher, A. (2018). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(8), 1269–1285.
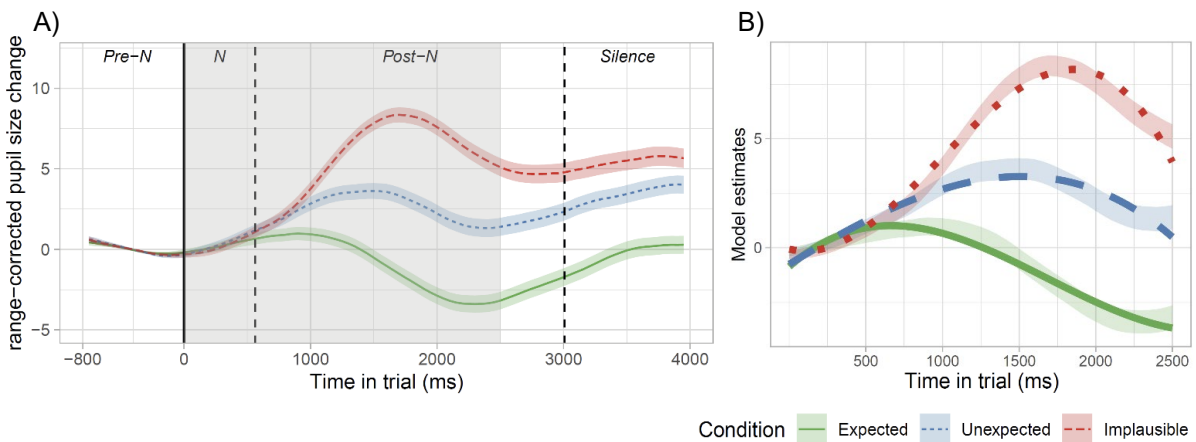[5] Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). *Journal of Cognitive Neuroscience*, 32(1), 12–35.

Figure 2. A) Change in pupil size in three conditions, time-locked to critical noun onset (0 ms). Semi-transparent shades represent standard error. The critical window for data analysis was marked in grey (0-2500 ms). B) Growth curve model fits overlaid on observed pupil size (standard error).
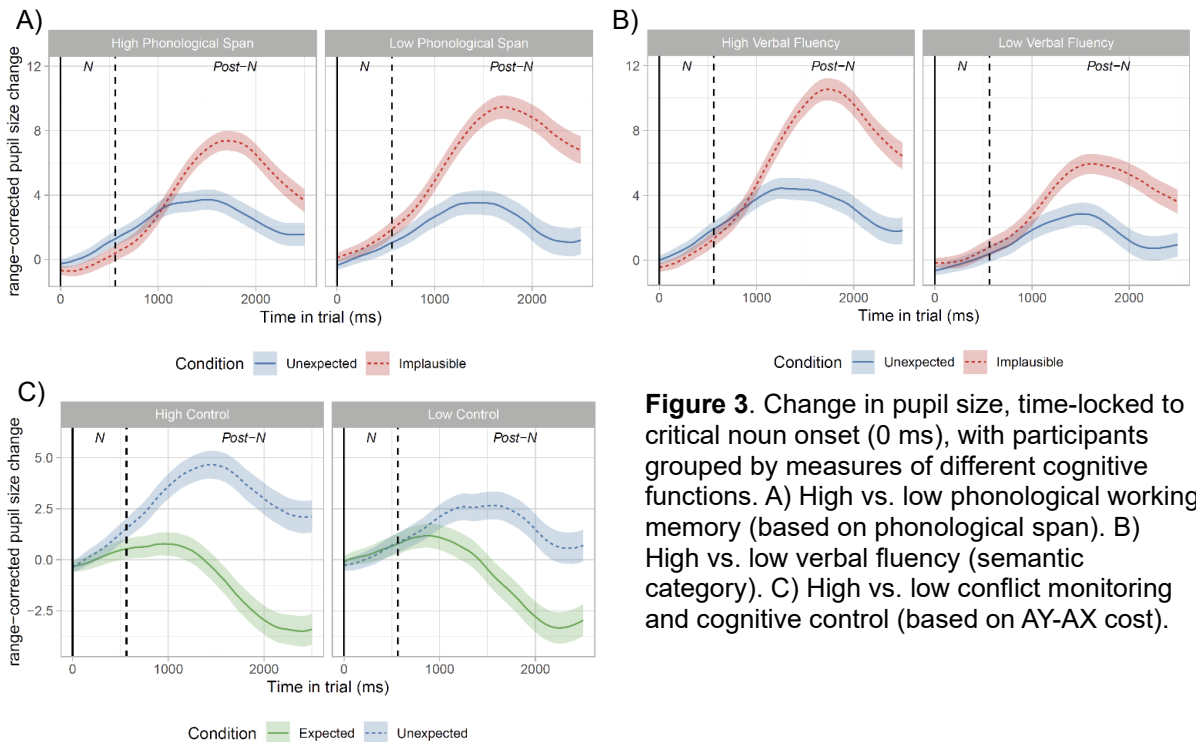

Figure 3. Change in pupil size, time-locked to critical noun onset (0 ms), with participants grouped by measures of different cognitive functions. A) High vs. low phonological working memory (based on phonological span). B) High vs. low verbal fluency (semantic category). C) High vs. low conflict monitoring and cognitive control (based on AY-AX cost).