

Title: Perceived interpretability predicts satiability for CNPC islands but not WH islands

Introduction. Recent studies report that the acceptability ratings for sentences tend to increase after repeated exposure (*the satiation effect*) [1-5]. Not all sentence types are affected by satiation, and those that satiate also demonstrate varying rates of satiation [1-6]. Previous work speculated that satiability is modulated by interpretability: more interpretable sentences tend to satiate faster [3,4]. The interpretability effect is predicted by the “adaptation account of satiation”, where satiation is analyzed as an increase in expectation for certain linguistic representations [6]. If no representation can be recovered from the strings (i.e., the sentences are uninterpretable), satiation should not occur. In the current study, we test the effect of two different forms of interpretability on satiability: *actual interpretability* (whether the comprehender gets the meaning as suggested by the sentence structure), and *perceived interpretability* (whether the comprehender believes they interpreted the sentence correctly).

Methods. We conducted an acceptability judgment experiment (N=400) testing sentences violating the complex-NP constraint (CNPC) and the *whether*-island constraint (examples in (1)), which have been shown to satiate [1-6], in addition to word salad and grammatical fillers. A total of 8 critical items were included. Participants answered a 4-alternative forced-choice comprehension question (example in (2)) and rated the acceptability of the target sentence on a continuous slider scale. Since recovering the filler-gap dependency is at the crux of interpreting island-violating sentences, the comprehension questions always probe the filler-gap association. Each question offers one correct choice, two incorrect distractors, and one “I don’t know” option. To increase the diversity of comprehension responses, each sentence was preceded by either a supportive context aiding the correct interpretation, or a neutral context. If *actual interpretability* modulates satiability, we expect sentences interpreted correctly to satiate faster than those with incorrect or “I don’t know” comprehension answers. If *perceived interpretability* modulates satiability, we expect sentences coupled with either correct or incorrect comprehension answers to show a greater rate of satiation than those with “I don’t know” answers, but no satiability distinction between those with correct and incorrect answers.

Results. Figures 1a,b show satiation plots for CNPC and *whether*-island sentences. There was no evidence that sentences with correct comprehension responses (representing *actual interpretation*) satiated faster than the rest (CNPC: $\beta=0.0012$, $SE=0.0009$, $t=1.33$, $p=0.19$; *whether*-island: $\beta=5e-05$, $SE=0.0009$, $t=0.05$, $p=0.96$). When correct and incorrect responses were grouped together to represent *perceived interpretation*, an LMER model predicting acceptability from order, perceived interpretability, and their interaction showed a significant interaction effect ($\beta=0.0028$, $SE=0.0012$, $t=2.18$, $p<0.05$) for CNPC sentences, but not for *whether*-island sentences ($\beta=0.0003$, $SE=0.002$, $t=0.12$, $p=0.91$), suggesting that perceived interpretation yields a greater satiation rate for CNPC. Finally, an LMER model predicting acceptability from trial order, context type, and their interaction showed no significant interaction ($\beta=0.0013$, $SE=0.0013$, $t=1.03$, $p=0.31$), suggesting that context type itself might not influence satiation rate. All models included the maximal random effect structure allowing convergence.

Discussion. This work suggests that *perceived interpretability* modulates satiability, at least for CNPC sentences. Moreover, supportive context increases the rate of *actual interpretation* of the target sentences, but does not increase the rate of satiation. The lack of interpretability effect for *whether*-island sentences is an interesting puzzle and deserves attention in future studies.

Reference [1] Snyder 2000 [2] Snyder (2022) [3] Francom (2009) [4] Hofmeister, Cassasanto and Sag (2013) [5] Lu, Frank and Degen (*under review*) [6] Lu, Lassiter and Degen (2021)

(1) Example of target stimuli:

CNPC: That's the bottle of water that the professor believes the claim that the student poured __ into a bowl of soup.

Whether-island: That's the bottle of water that the professor wonders whether the student poured __ into a bowl of soup.

(2) Example of comprehension questions

Q: The student might have poured _____.

A. A bottle of water B. A bowl of soup C. A bowl of water D. I don't know

(3) Example of contexts

Supportive context: The professor has a student. Someone claimed that the student poured something.

Neutral context: The professor has a student. Someone claimed that the student was in a hurry.

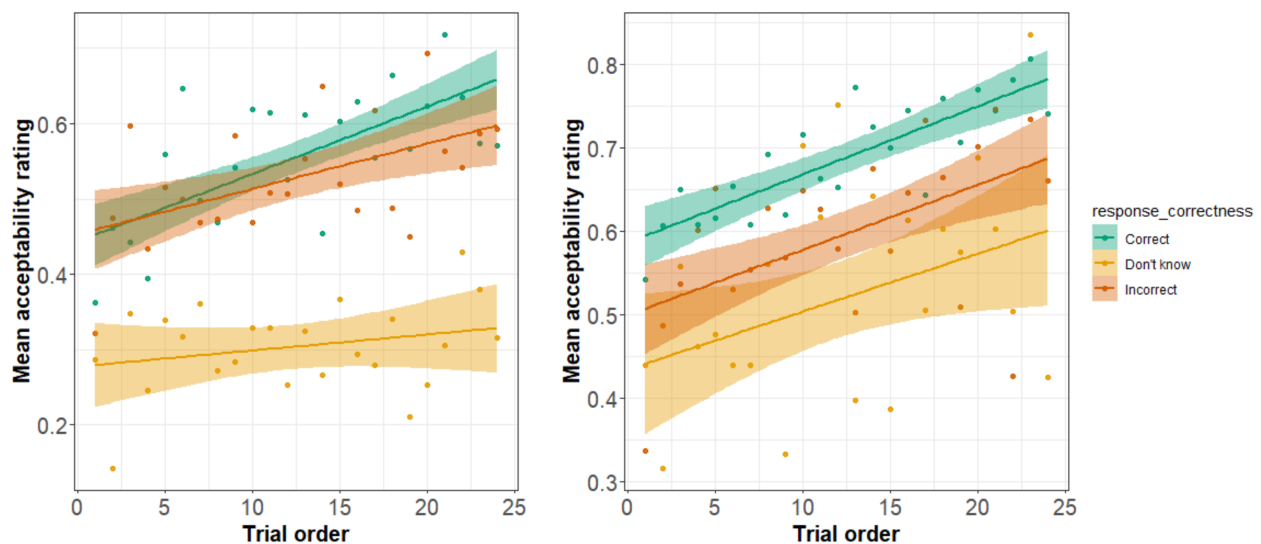


Fig.1a By-response satiation plot for CNPC Fig.1b By-response satiation plot for *whether*-island