

Trial Item Content Impacts Island Effects in Online Processing of English
Dennis Ryan Storoshenko and Jesse Weir - University of Calgary

Issue In a 48-participant **pilot study**, and a 136-participant **preliminary study**, we investigate the role of lexical content on seemingly unrelated grammatical processing, using *wh*-extraction across a weak island as a test case. Using self-paced reading (SPR), we examine whether manipulating semantic (im)plausibility via the relationship between extracted *wh*-phrase’s content and adjacent lexical material impacts processing at the point an island is overtly signalled. Further, we examine whether effects persist into the island, between items that are not part of the *wh*-chain. In items whose (im)plausibility lies in evoking gender stereotypes, we answer yes to both questions. We make no strong claim on the nature or origin of this interaction between gender stereotype processing and islands, but we reject the null hypothesis that there is no such interaction.

Background Prior SPR studies have shown readers to slow down at the point entering an embedded clause where an extraction island is signalled, before encountering the *wh*-gap (Moscati, 2014). Further, research in other languages has shown that gender mismatches interfere with reference resolution (Xu et al., 2013) and grammatical processing (Molinaro et al., 2016). We thus hypothesize that either confirming or subverting a gender stereotype may impact reading times at a weak island boundary and at a later pronoun, based on the Phillips (2006) evidence that participants are sensitive to lexical content when filling gaps.

Study Design The table below shows a paradigm of stereotyped test items in our SPR task. The “stereotype” region contains two variable manipulations. The **bias (Fem/Masc)** of stereotype, along with whether the *wh*-extracted item **(mis)matches the stereotype**, is evident at region 3, setting the first two variables. At region 9, the “pronoun” region, a gendered **pronoun (mis)matches** the stereotyped matrix subject, resulting in a 2x2x2 design. 16 trial items were constructed and presented in a Latin-Square design. Eight neutral items were also presented, with non-gendered professions at the stereotype region. In these, semantic congruence is manipulated at the stereotype region, while the pronoun region alternates between an even balance of *he/she* vs. *they* trials. For the stereotyped and neutral items, region 6, the “island” region, contains a *wh*-word establishing a weak island. 26 additional distractor items, grammatical and ungrammatical, were presented. Data were collected online using PsychoPy 3. Residual reading times (Trueswell et al., 1994) at all critical regions and their spillovers are analyzed using lme4 and lmerTest in R.

M-Match-Match	1 Which tie - 2 did - 3 the general - 4 forget - 5 completely - 6 where - 7 after - 8 the parade - 9 he - 10 had - 11 accidentally - 12 left - 13 it?
M-Match-Mismatch	1 Which tie - 2 did - 3 the general - 4 forget - 5 completely - 6 where - 7 after - 8 the parade - 9 she - 10 had - 11 accidentally - 12 left - 13 it?
M-Mismatch-Match	1 Which lipstick - 2 did - 3 the general - 4 forget - 5 completely - 6 where - 7 after - 8 the parade - 9 he - 10 had - 11 accidentally - 12 left - 13 it?
M-Mismatch-Mismatch	1 Which lipstick - 2 did - 3 the general - 4 forget - 5 completely - 6 where - 7 after - 8 the parade - 9 she - 10 had - 11 accidentally - 12 left - 13 it?

Results In our pilot study, comparing eight grammatical long-distance *wh*-extractions with complementizer *that* at region 6 against the neutral items reveals a significant slowdown at the island region ($p < 0.01$), confirming sensitivity to the island. In the preliminary study, neutral items show no significant effects at the stereotype or island regions, nor their spillovers. At the pronoun region, *they* is read significantly faster ($p = 0.04$), but this falls away by the spillover region 10. In trial items, the stereotype region shows no effect, but at its spillover region 4, masculine items are read significantly faster ($p = 0.04$). Regardless of (mis)matching, this effect persists, marginally at the island region ($p = 0.11$), and stronger at the spillover region 7 ($p = 0.02$). Trial items show no effects at the pronoun region, but a marginal slowdown for mismatching pronouns emerges at region 10; model comparisons show a fixed factor of pronoun (mis)match still yields the best fit for the data.

Discussion The sustained stereotype effect in trial items presents a warning to researchers: such effects can impact participant response times at critical processing regions seemingly unrelated to the stereotyped content. A **replication study** is planned before June, to establish whether the observed effects are significant across equal-sized participant pools.

References

- Molinaro, Nicola, Jui-Ju Su, and Manuel Carreiras. 2016. Stereotypes override grammar: Social knowledge in sentence comprehension. *Brain & Language* 155-156:36–43.
- Moscato, Vincenzo. 2014. Arguments, adjuncts, and lexical restrictions: Reaction-time evidence for a selective intervention effect in wh-questions. *Journal of Advanced Linguistic Studies* 3:232–248.
- Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82:795–823.
- Trueswell, John C., Michael K. Tanenhaus, and Susan M. Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33:285–318.
- Xu, Xiaodong, Xiaoming Jiang, and Xiaolin Zhou. 2013. Processing biological gender and number information during Chinese pronoun resolution: ERP evidence for functional differentiation. *Brain & Cognition* 81:223–236.