

Timing and (mis)interpretation of NPI illusions

Hanna Muller (University of Maryland), Celeste Joly (University of Maryland), Iria de Dios Flores (Universidade de Santiago de Compostela and Basque Center on Cognition, Brain, and Language), Philip Resnik (University of Maryland), and Colin Phillips (University of Maryland)
hmuller@umd.edu

Negative polarity illusions were once taken as a prime example of the key role of memory retrieval operations in sentence processing [4]. Negative polarity items (NPIs) are words and phrases such as *ever* and *any*, which are only acceptable (“licensed”) in negative (or similar) contexts (ex. 1 vs 4). Sentences which contain an NPI and a nearby but irrelevant negative context (2) are accepted more often than other ungrammatical sentences such as (4) [2,5]. Early explanations focused on computations that occur at the NPI, namely a retrospective search for an appropriate licenser [4]. However, recent work has shown that the illusion profile is more restricted than this account predicts [1,3], and suggests that the representation of material prior to the NPI plays an important role. Here we show the division of labor between early interpretive processes and processes that occur directly at the NPI in creating NPI illusions. We consider two hypotheses for how prior material could be (mis)represented – which focus, respectively, on the scope of the negative quantifier and on the recency of the relative clause (RC) meaning – and demonstrate in two experiments that only the second of these captures the profile of the illusion.

The first explanation we consider is that negative quantifiers are occasionally erroneously interpreted as if they take wide scope over the entire sentence (including the NPI). This hypothesis successfully accounts for the finding that illusions occur when the intrusive negative word is a quantifier but not when it’s simple sentential negation as in (3) [1]. It additionally predicts that the sentences that give rise to illusions are interpreted as globally negative before the NPI is seen. To test this prediction, we conducted an interpretation study (MTurk; N=32). Participants were presented with both sentences like (1)-(4) and equivalent sentences without NPIs (6)-(9) allowing us to test interpretations that arise as a result of the position and type of negation, which could have been established prior to the NPI in illusion sentences. After the sentence, participants were asked whether the sentence was acceptable (yes/no) (in order to filter data to only trials that were accepted) and answered a comprehension question like (10) (yes/no/I don’t know). In illusion-type sentences without an NPI, globally negative interpretations are reported in very few trials, suggesting that QR of the negative quantifier is not the cause of NPI illusions. However, in illusion-type sentences that include an NPI the rate of globally negative interpretations is sharply higher, suggesting that key interpretive processes occur when the NPI appears.

Turning to our second approach, we suggest that the previously established contrast in licenser types (*no* vs. *haven’t*) can be explained as an effect of representation of RC meaning. That is, negative quantifiers like *no* may be used primarily to make strong negative claims where NPIs may be expected, whereas simple sentential negation is compatible with a broader range of speaker meanings. Under this hypothesis, NPI illusions occur because the comprehender mistakes a nearby semantically compatible negative RC for the NPI’s local environment. Because the focus here is on the meaning of the entire RC, this hypothesis predicts that NPIs farther away from the RC edge will be less prone to illusions. In experiment 2 we show that the distance effect found in prior work [3] is in fact about the distance from the RC (B in 5) and not (as the original authors state) about the distance from the licenser (A+B in 5). We conducted a speeded acceptability study (MTurk; N=37) and find that added material inside the RC has no impact on the magnitude of the illusion relative to short-distance controls, but added material *outside* the RC leads to the disappearance of the illusion. These data are consistent with the hypothesis that illusions arise because of the compatibility between the nearby RC meaning and the NPI.

Thus we demonstrate that illusions are sensitive to the proximity of the relative clause and do not require that globally negative interpretations be established prior to the NPI. We suggest that NPI illusions reflect an attempt to reconcile the conflict between an accurate syntactic parse with an impression of acceptability caused by the nearby RC meaning.

(1) **No** authors [that the critics have recommended for the award] have **ever** received acknowledgement for a best-selling novel.

(2) The authors [that **no** critics have recommended for the award] have **ever** received...

(3) The authors [that the critics **haven't** recommended for the award] have **ever** received...

(4) The authors [that the critics have recommended for the award] have **ever** received...

(5) The authors [that **no** critics have recommended for the award] have received **any**



acknowledgement for a best-selling novel.

(6) **No** authors [that the critics have recommended for the award] have received acknowledgement for a best-selling novel.

(7) The authors [that **no** critics have recommended for the award] have received...

(8) The authors [that the critics **haven't** recommended for the award] have received...

(9) The authors [that the critics have recommended for the award] have received...

(10) Have the authors received acknowledgement for a novel? Yes / No / I don't know

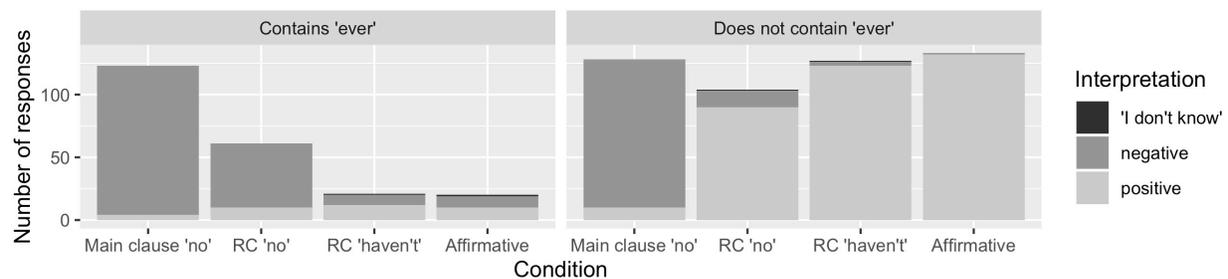


Figure 1: Only trials that were judged acceptable are plotted. Illusion sentences (RC 'no', sentence contains 'ever') are interpreted as expressing a globally negative claim 84% of the time, whereas equivalent sentences without 'ever' are interpreted as expressing a globally negative claim 13% of the time, suggesting that wide-scope interpretations of quantifiers do not cause NPI illusions.

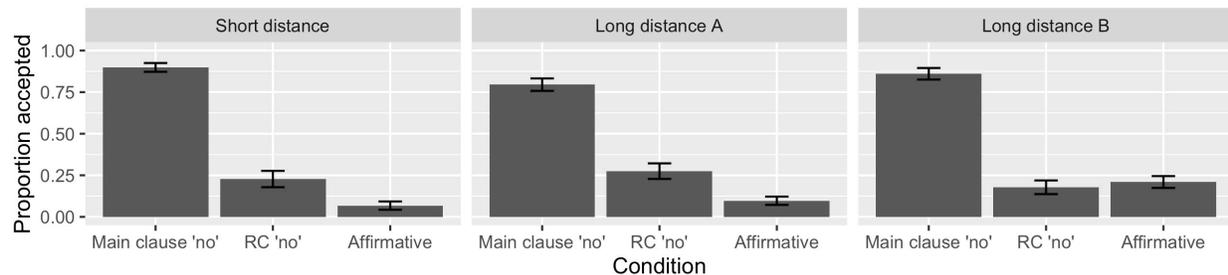


Figure 2: Short distance conditions replicate the standard illusion effect; Long distance A conditions show that the illusion persists when additional material intervenes between the negative word and the NPI (span A in sentence 5); Long distance B conditions replicate Parker & Phillips's experiment 2, showing that the illusion disappears when additional material intervenes between the relative clause edge and the NPI (span B in sentence 5).

References

- [1] de Dios Flores, Muller, & Phillips (2017) Negative polarity illusions: licensors that don't cause illusions, and blockers that do. Poster at 30th Annual CUNY Conference on Human Sentence Processing, MIT. [2] Drenhaus, Saddy, & Frisch (2005). Processing negative polarity items: When negation comes through the backdoor. *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 145-165 [3] Parker & Phillips (2016). Negative polarity illusions and the format of hierarchical encodings in memory. *Cognition*, 157, 321-339. [4] Vasissth, Brüssow, Lewis, & Drenhaus (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32, 685-712. [5] Xiang, Dillon & Phillips (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, 108, 40-55.

Negative polarity illusions were once taken as a prime example of the role of memory retrieval operations in sentence processing. However, recent work shows that the illusion profile is more restricted than this account predicts. We show the division of labor between early interpretive processes and processes that occur at the NPI in creating NPI illusions. We consider two hypotheses for how prior material could be (mis)represented –which focus, respectively, on the scope of negative quantifiers and on the recency of the relative clause meaning –and demonstrate in two experiments that only the second captures the profile of the illusion.