

Adults demonstrate implicit causality and consequentiality biases for novel verbs

Amanda Rose Yuile¹ (ayuile@mghihp.edu) & Joshua K. Hartshorne¹

1. MGH Institute of Health Professions

In causally dependent clauses (1) comprehenders tend to interpret subject pronouns as referring to the cause entailed by the verb (*implicit causality*). Similarly, when the second clause describes a consequence of the first (2), comprehenders tend to interpret subject pronouns as referring to the affected entity (*implicit consequentiality*). We refer to these biases as *IC biases*.

(1) Archibald angered Bartholomew because he...

(2) Because Archibald angered Bartholomew, he...

IC biases of individual verbs are strongly correlated with the verb's argument structure, i.e. how event participants (e.g., AGENT or PATIENT) are mapped to grammatical positions in sentences (e.g., SUBJECT or OBJECT) [1-3]. This is because verb argument structure is a systematic function of the verb's meaning [4].

While the existence of IC biases is well-attested [1-3,5-6], researchers disagree about the origin. IC biases may emerge by tracking probabilistic patterns of reference in the input [7], or may emerge for free from representations of verb meaning and discourse structure and, therefore, are available as soon as the relevant linguistic knowledge is acquired [1-2]. Thus, competing theoretical accounts of IC biases propose different roles of experience in learning.

To better understand the role of experience in the acquisition of IC pronoun biases, we asked whether adults extend IC biases to transitive emotion verbs they have never encountered before. (Transitive emotion verbs make up a significant portion of the verbs known to result in IC pronoun biases [1-2,5].) We specifically asked whether verb argument structure judgments predicted participants' pronoun resolution patterns. A positive relationship between verb argument structure and pronoun interpretation would indicate generalization of IC biases to novel psych verbs. A lack of a relationship would suggest item-based learning.

We built from prior work [7] showing that changing the meaning of a verb can reliably affect beliefs about the verb's argument structure. Thus, we generated 28 unique definitions for novel transitive emotion verbs. (A separate norming study ensured the novel verb definitions did not have an obvious English translation.) On each trial, participants were first presented with a novel verb and definition: *dax -- the love of all things good*. Then, we confirmed whether they interpreted the verb as assigning the experiencer to the subject position or object position:

Suppose Albert felt dax in relation to Bernie. Which of the following sentences is the best way to describe that? (a) Albert daxed Bernie. (b) Bernie daxed Albert.

Participants choosing (a) mapped experience to subject; those choosing (b) mapped it to object.

As expected, participants were far more likely to judge the verbs to be experiencer-subject for some definitions than for others [see 8]. This allowed us to then see whether the argument structure they assigned affected their subsequent pronoun interpretation. Participants were asked to indicate their pronoun interpretation preferences for the verb in either implicit causality sentences (Experiment-1a: *John daxed Sarah because he / she*) or implicit consequentiality sentences (Experiment-1b: *Because John daxed Sarah, he / she*). In Experiment-1a, 688 English-speaking participants each judged 6 verbs; in Experiment-1b, 97 English-speaking participants each judged 12 verbs. Logistic regression showed that verb argument structure (as assessed by participants) was significantly predictive of pronoun interpretations in Experiment-1a ($p = .04$) and Experiment-1b ($p < .001$). See Figures 1-4.

In sum, results show that verb argument structure was not selected at random, but systematically varied by the definition given for the verb and this in turn influenced pronoun interpretation. These findings offer strong constraints on the role of experience in learning IC biases, suggesting that these biases are not computed over individual lexical items but reflect more abstract generalizations.

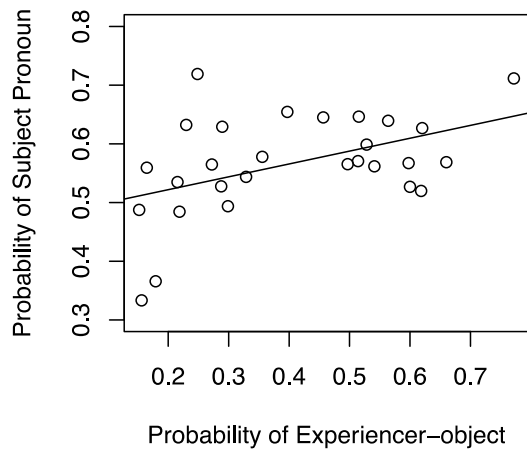


Figure 1. Correlation between argument structure judgments and subject pronoun resolution in implicit causality sentences.

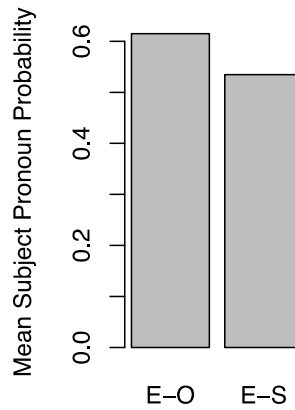


Figure 2. Average probability of subject pronoun by verb argument structure in implicit causality sentences.

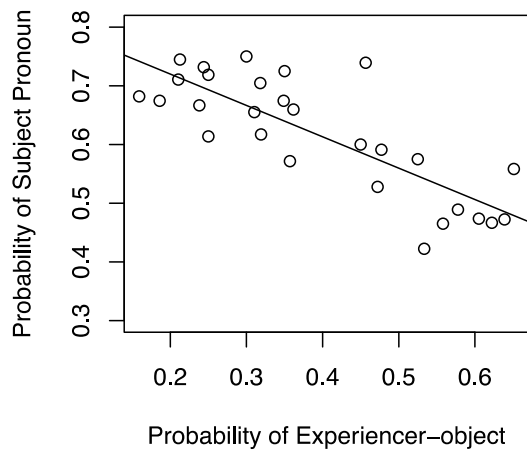


Figure 3. Correlation between argument structure judgments and subject pronoun resolution in implicit consequentiality sentences.

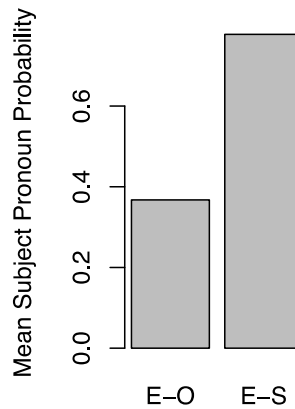


Figure 4. Average probability of subject pronoun by verb argument structure in implicit consequentiality sentences.

References

1. Hartshorne, J.K. & Snedeker, J. (2013). *Language and Cognitive Processes*.
2. Hartshorne, J.K. et al. (2015). *Language, Cognition and Neuroscience*.
3. Hartshorne, J.K., Sudo, Y., & Urwashi, M. (2013). *Experimental Psychology*.
4. Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*.
5. Ferstl, E.C. et al. (2011). *Behavior Research Methods*.
6. Garvey, C. & Caramazza, A. (1974). *Linguistic Inquiry*.
7. Guan, S. & Arnold, J.E. (2021). *Discourse Processes*.
8. Harthorne et al. (2016). *Cognition*.