

Grammatical and Coherence-Driven Biases in Pronoun Interpretation

In contrast to theories that argue that pronoun interpretation is guided largely by grammatical factors, Hobbs (1979) famously argued that pronoun interpretation is instead simply a by-product of more general logical inferencing involved in establishing COHERENCE RELATIONS in discourse. Unlike the grammatical-factor approach, however, the Hobbsian view does not explain results from Stevenson et al. (1994), who found an asymmetry between pronoun production and interpretation in passage completion studies. For instance, when writing completions to Source-Goal transfer-of-possession passages with a pronoun prompt (1a), participants were equally likely to interpret the pronoun to refer to the non-subject Goal (*Bill* in 1a) as to the subject-position Source (*John*). However, in passages with a full-stop prompt (1b), participants were far more likely to produce a pronoun when they re-mentioned the subject/Source, whereas they were far more likely to produce a name when they re-mentioned the non-subject/Goal (see also Arnold, 2001). Because pronouns are simply represented as unbound variables in Hobbs' system, no misalignment between interpretation and production biases is expected.

- (1) a. John passed the comic to Bill. He _____
b. John passed the comic to Bill. _____

Rohde et al. (2006, 2007) followed-up Stevenson et al.'s study by showing that the even distribution of pronoun interpretations in the pronoun-prompt condition masks a much stronger system of biases that emerge when passages are categorized by coherence relation. For instance, whereas passages with OCCASION relations exhibited a strong Goal bias, relations such as EXPLANATION and ELABORATION exhibited a strong Source bias, thereby confirming the role of coherence establishment in pronoun interpretation. Left unresolved, however, is how this evidence for the Hobbsian view can be reconciled with the evidence for a grammatical-role bias in pronoun production.

Indeed, the existence of a subject bias for pronouns is predicted to have repercussions for coherence establishment: Because the mere presence of a pronoun – even one whose reference is ambiguous as in (1a) – will bias interpretation towards the subject of the last clause, encountering a pronoun would be expected to cause a shift in comprehenders' expectations in favor of Source-biased coherence relations. To test this prediction, we had participants complete passages in a pronoun-prompt condition (1a) and in a full-stop condition (1b). Judges categorized the resulting passages by coherence relation and annotated the intended referents of all referring expressions. As predicted, the distribution of coherence relations between the two conditions differed: There were significantly more instances of Source-biased coherence relations in the pronoun-prompt condition.

These results demonstrate not only that coherence establishment biases influence pronoun interpretation expectancies, but conversely that independent biases in pronoun production influence expectations about ensuing coherence relations. A Bayesian model of pronoun interpretation ($p(\textit{referent}|\textit{pronoun})$) – in which factors (e.g., subjecthood) that determine whether a speaker will choose a pronoun instead of a name ($p(\textit{pronoun}|\textit{referent})$) can be distinguished from factors (e.g., coherence relations) that generate expectations about who will be mentioned next ($p(\textit{referent})$) – captures our data and explains Stevenson et al.'s interpretation/production asymmetry.