

# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

### Title

Listener Knowledge Structures Commonsense Explanation

### Permalink

<https://escholarship.org/uc/item/66p2499n>

### Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

### Authors

Machino, Yuka

Shprints, Ron

Siegel, Max

et al.

### Publication Date

2024

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Listener Knowledge Structures Commonsense Explanation

**Yuka Machino**

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

**Ron Shprints**

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

**Max Siegel**

MIT, Cambridge, Massachusetts, United States

**Lionel Wong**

Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

**Josh Tenenbaum**

MIT, Cambridge, Massachusetts, United States

## Abstract

We tailor the explanations we give depending on the person asking for them – you would explain why an event happened differently depending on which of the contributing causes the listener already knows. While significant prior work focuses on how causal structure in the world influences explanation, we focus on how explanation production is modulated by listener belief. We propose a computational model framing explanation as rational communication about causal events, using a recursive theory-of-mind and language production framework to choose amongst possible explanatory utterances that minimize the divergence between speaker and listener belief about a why an event happened. We evaluate our model using some partial observer stimuli, which manipulate the listener’s stated prior knowledge about an event, and find that our model well-predicts human judgements about which of several contributing causes is the best explanation for a speaker to provide by modeling their communicative value to the listener.