NEURIPS 2026 · POSITION PAPER TRACK

Anthropomorphic Language in AI Discourse Can Be Constructive

Yingshan Chang · Yonatan Bisk

Carnegie Mellon University

OVERVIEW

Anthropomorphic language attributes human traits to non-human entities

Criticisms:

Blur human–machine differences

Perpetuate AI hype

Promote unfair agendas such as solutionism and transhumanism

Instead of adding to the criticism,

we ask: why do our tendency and resistance to anthropomorphizing AI occur?

Then we argue that it can be constructive.

CENTRAL ARGUMENT

Anthropomorphic language in AI discourse can be constructive ---

by serving as a provisional language that illuminates unresolved research landscape.

We support our position by advancing three novel claims:

1

Descriptive

It results from ordinary mechanisms of language use.

2

Normative

Suppressing it is worse than treating it constructively.

3

Interventionist

Interventions can productively use it as a signal (Sections 4–6).

CLAIM 1 · DESCRIPTIVE

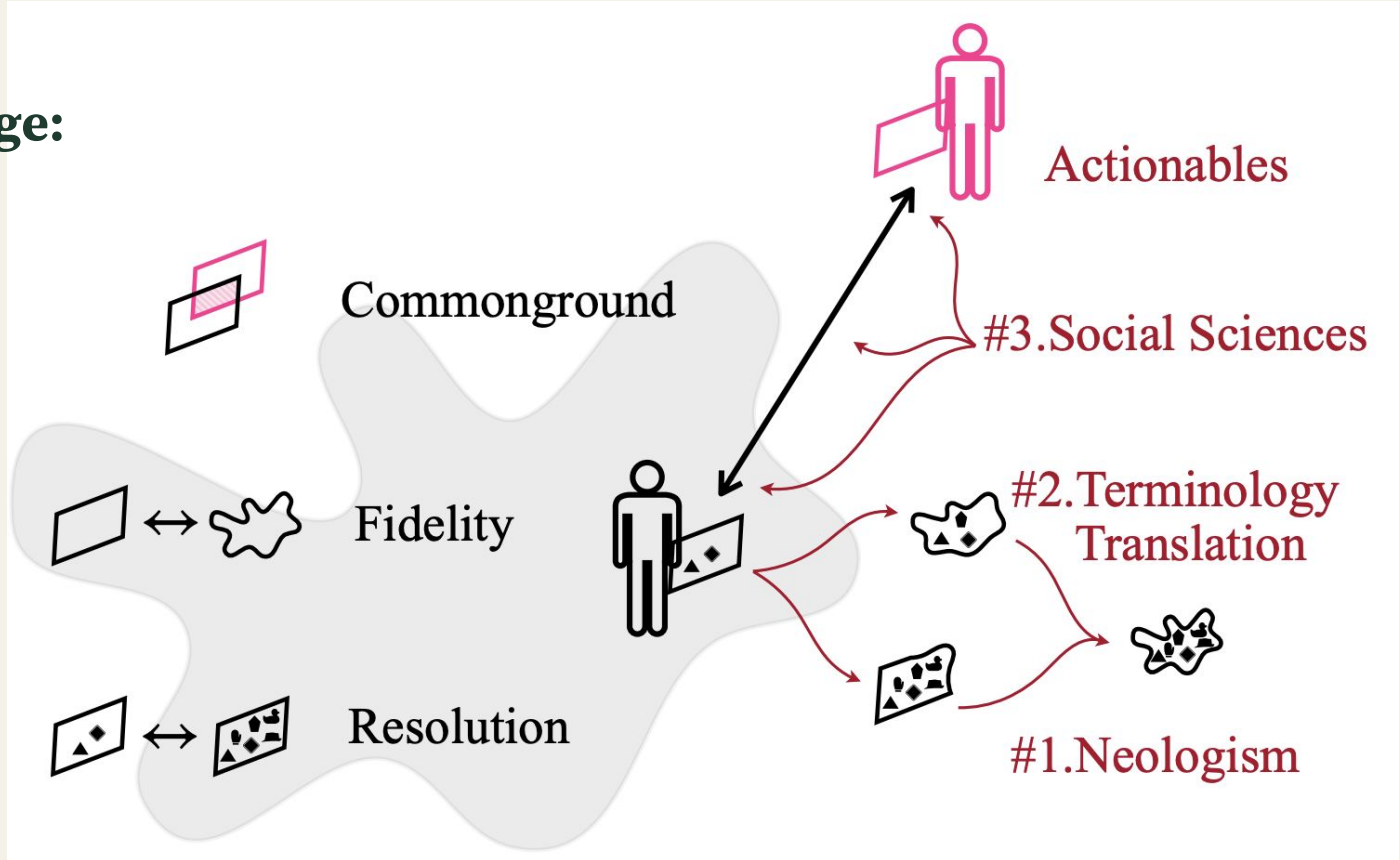
Anthropomorphic language results from ordinary mechanisms of language-use.

What gives rise to anthropomorphic language:

Unshared vocabulary

Imperfect vocabulary

Limited vocabulary

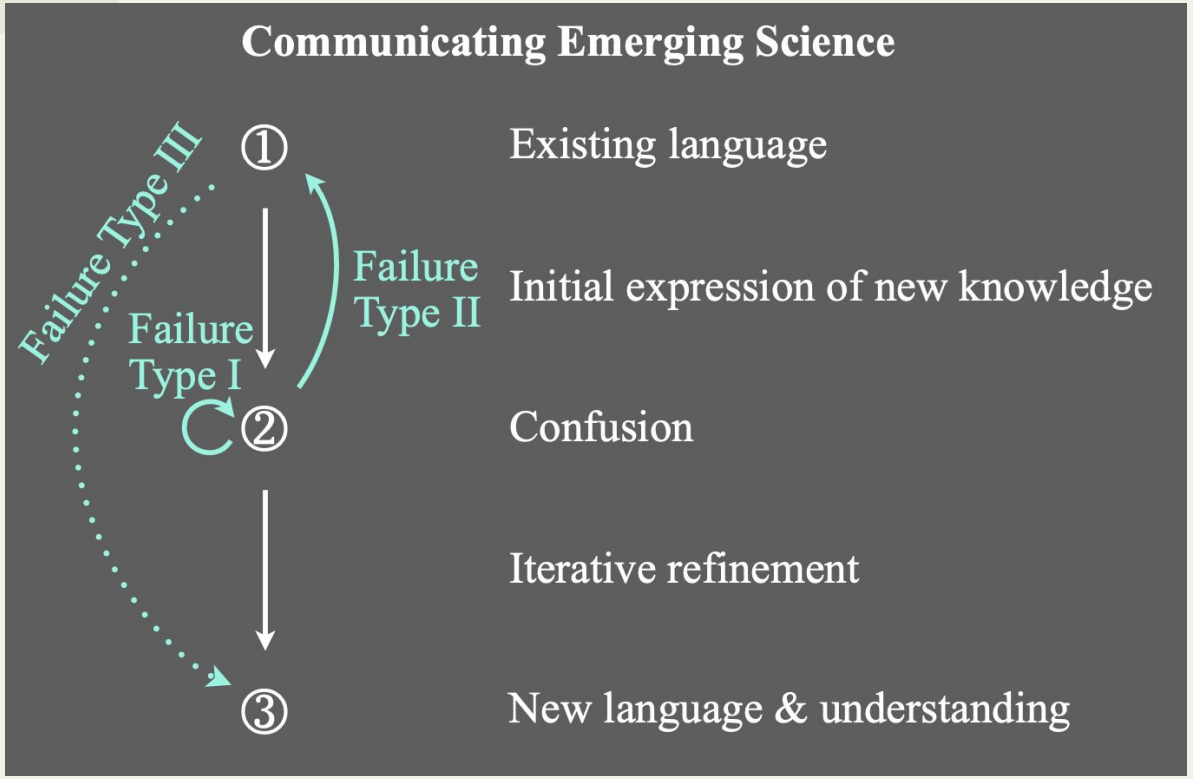


CLAIM 2 · NORMATIVE

Suppressing anthropomorphic language is *worse* than treating it constructively.

Initial adequacy --> Confusion --> Fresh clarity

AI terminology is here



THREE WAYS TO GET IT WRONG

Premature satisfaction

Conservatism

Idealization

CLAIM 3 · INTERVENTIONIST

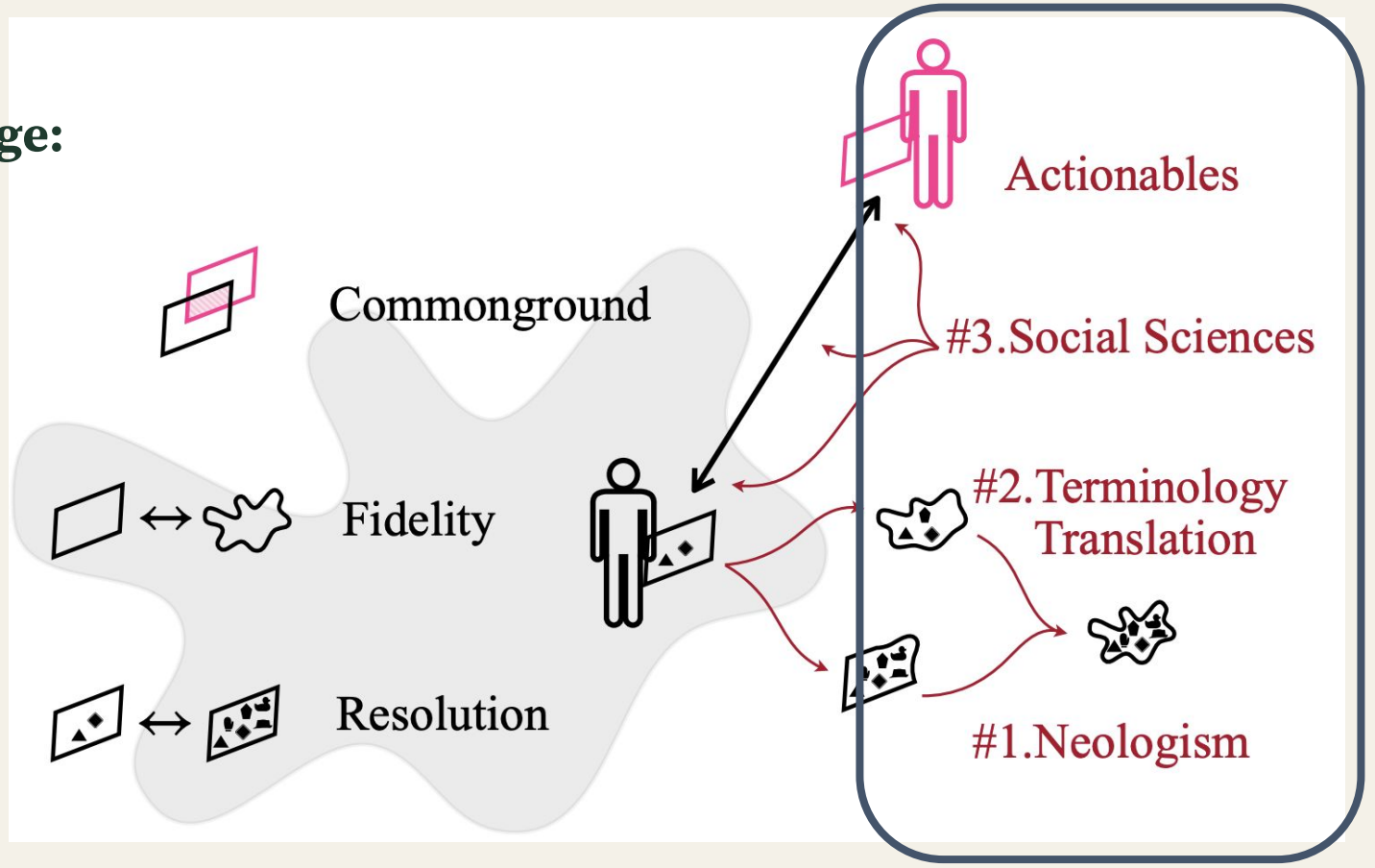
Various interventions can productively use anthropomorphic language as a signal.

What gives rise to anthropomorphic language:

Unshared vocabulary

Imperfect vocabulary

Limited vocabulary



HIGHLIGHTS

Discussion Points

Conditional

we do not mean that anthropomorphic language is inherently or automatically constructive.

Falsifiable

Our position can be weakened if further evidence show that:
(i) anthropomorphic language chiefly stems from poor understanding or careless word choice
(ii) minimizing such language reliably improves cross-disciplinary understanding and public trust
(iii) interventionist strategies that we propose do not in fact reduce confusion or improve clarity

Reciprocity between *saying* and *doing*

Scientific language and terminology shapes the research agenda itself.

NEURIPS 2026 · POSITION PAPER TRACK

Thank You!