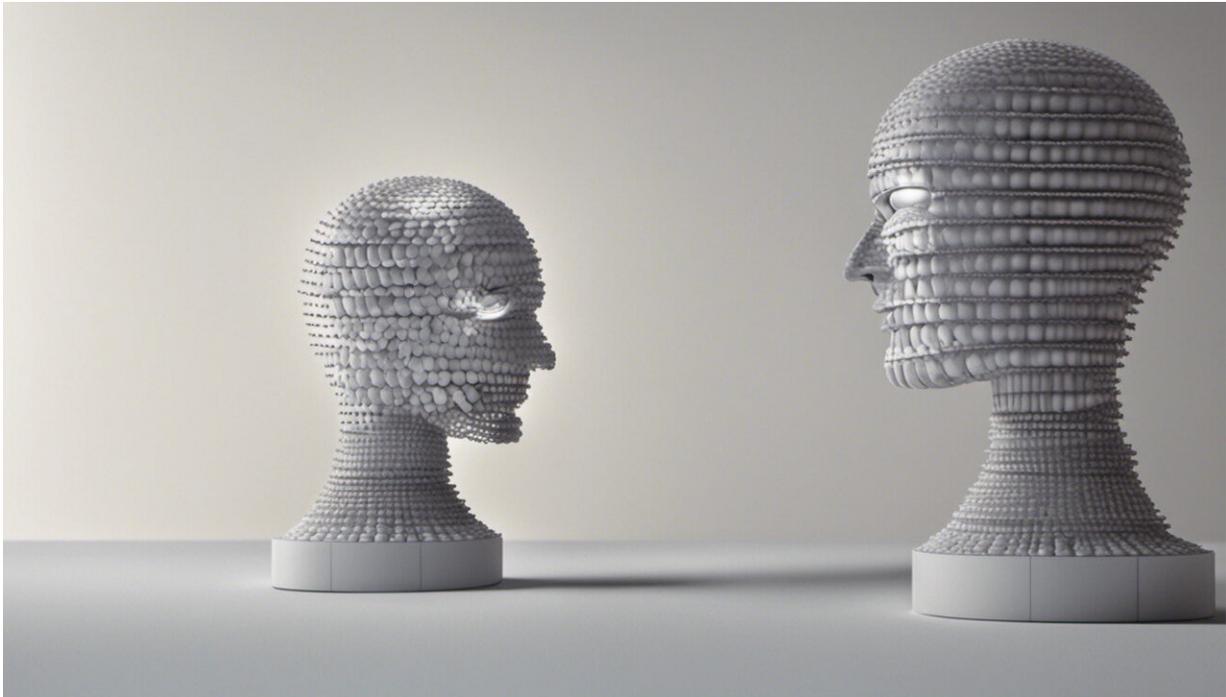


Teaching chatbots how to do the right thing

March 29 2018, by Amanda Turnbull



Credit: AI-generated image ([disclaimer](#))

In this age of information —and misinformation —advancements in technology are challenging us to rethink how language works.

Take conversational chatbots, for example. These computer programs mimic human conversation via text or audio. The mattress company Casper created [Insomnobot-3000](#) to communicate with people who have sleep disorders. It gives those who have trouble sleeping the opportunity

to talk to "someone" while everyone else is asleep.

But Insomnobot-3000 doesn't only chit-chat with its users, answering questions. It aims to reduce the loneliness felt by sufferers of insomnia. Its words have the potential to have an impact on the human user.

At its most basic, language does things with words. It is a form of action that does more than simply state facts.

This fairly straightforward observation was made in the 1950s by an obscure and slightly eccentric Oxford University philosopher, [John Langshaw Austin](#). In his book, *How To Do Things With Words*, Austin developed the concept of performative language.

What Austin meant was that language doesn't just describe things, it actually "performs." For example, if I say I bequeath my grandmother's pearl necklace to my daughter, I am doing more than simply describing or reporting something. I am making a meaningful action.

Austin also classified speech into three parts: Meaning, use and impact. His study and findings on language became known as speech-act theory. This theory was important not only in philosophy, but also in other areas such as law, literature and feminist thought.

A prescription for the chatbot industry

With this in mind, what can Austin's theory tell us about today's conversational chatbots?

My research focuses on the intersection of law and language, and what Austin's theory has to say about our understanding of how creative machinery is changing traditional societal operations, such as AI writing novels, robo-reporters penning news articles, [massive open online](#)

[courses](#) (MOOCs) replacing classrooms and professors using essay-grading software.

[Current chatbot technology](#) is focused on improving chatbots' ability to mimic the meaning and use of speech. A good example of this is [Cleverbot](#).

But the chatbot industry should be focused on the third aspect of Austin's theory —determining the impact of the chatbot's speech on the person using it.

Surely, if we are able to teach chatbots to mimic the meaning and use of human speech, we should also be able to teach them to imitate its impact?

Learning to have a conversation

The latest chatbots rely on cutting-edge machine learning, known as deep learning.

[Machine learning](#) is an application of AI that can learn without human help. Deep learning, which is modelled after the network of neurons in the human brain, takes machine learning even farther. Data is fed into deep artificial neural networks that are designed to mimic human decision-making.

Chatbots designed with this [neural network technology](#) don't just parrot what is said or produce canned responses. Instead, they learn how to have a conversation.

Chatbots analyze massive quantities of human speech, and then make decisions on how to reply after assessing and ranking how well the possibilities mirror that speech. Yet despite these improvements, [these](#)

[new bots still suffer from the occasional faux pas](#) since they concentrate mainly on the meaning and use of their speech.

Earlier chatbots were far worse. In less than 24 hours of being released on Twitter in 2016, Microsoft's chatbot, an [AI system called Tay](#) (an abbreviation formed from "Thinking About You") and modelled after a teenage girl's language patterns, had more than [50,000 followers and had produced over 100,000 tweets](#).

As Tay greeted the world, her first tweets were innocent enough. But then she began to imitate her followers.

She quickly became a racist, sexist and downright distasteful chatbot. [Microsoft was forced to take her offline](#).

Tay had been entirely dependent on the data being fed to her —and, more importantly, on the people who were making and shaping that data. She did not understand what the human users were "doing" with language. Nor did she understand the effects of her speech.

Teaching chatbots the wrong thing

Some researchers believe that the more data chatbots acquire, the less offence they will cause.

But accounting for all possible responses to a given question could take a long time or rely on a lot of computing power. Plus, this solution of gathering more data on meaning and use is really just history repeating itself. [Microsoft's "Zo,"](#) a successor to Tay, still struggles with difficult questions about politics.

Put simply, the chatbot-industry is heading in the wrong direction —the chatbot-industry is teaching chatbots the wrong thing.

Transformative chatbots

A better chatbot would not only look at the meaning and use of words, but also the consequences of what it says.

Speech also functions as a form of social action. In her book *Gender Trouble*, philosopher [Judith Butler](#) looked at the performativity of language and how it heightens our understanding of gender. She saw gender as something one does, rather than something one is —that it is constructed through everyday speech and gestures.

Conversational chatbots are intended for diverse audiences. Focusing on the effect of speech could improve communication since the chatbot would also be concerned with the impact of its words.

In a tech industry challenged by its [lack of diversity](#) and inclusivity, such a chatbot could be transformative, as Butler has shown us in the construction of gender.

There is, of course, a caveat. Focusing on the impact of language is the defining trait of hoaxes, propaganda and misinformation —"fake news" —a deliberately engineered speech act, concerned only with achieving effect. No matter its form, fake news merely mimics journalism and is created only to achieve an effect.

Austin's theory of performativity in [language](#) helped us figure out how to speak to one another.

The chatbot industry should concentrate its efforts now on the impact of speech, in addition to the work already done on the meaning and use of words. For a [chatbot](#) can only be truly conversational if it engages in all aspects of a [speech](#) act.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: Teaching chatbots how to do the right thing (2018, March 29) retrieved 19 April 2024 from <https://techxplore.com/news/2018-03-chatbots.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.