Responsibility for Recommendations
UCSC, 15 February 2024
Mike Barnes, Australian National University

0. **Introduction**:

The Problem of Platform-Enabled Harms:
> "In 2017, the Rohingya were killed, tortured, raped, and displaced in the thousands as part of the Myanmar security forces' campaign of ethnic cleansing. In the months and years leading up to the atrocities, Facebook's algorithms were intensifying a storm of hatred against the Rohingya which contributed to real-world violence."
> - Agnès Callamard, *Amnesty International*

Against this backdrop, questions of *online intermediary responsibility* are currently unsettled. Legal liability is approached differently in different jurisdictions, and the underlying (moral) questions are marred by difficulties in conceptualizing what role platforms play in our speech.

My Claims: This talk argues social media companies share in the responsibility for the harms that occur on their platforms by demonstrating how they are *constitutive intermediaries*. I use *speech act theory* to illuminate the contribution that platform companies make to our communicative acts, showing how platforms shape users' speech, and also perform speech acts themselves. I examine three key features of online environments: *algorithms, affordances, amplification*, and in each case show how speech act theory aids our understanding of the ontology of speech acts on social media, enabling a more accurate attribution of responsibility when platforms are used to harm.

Three observations from Speech Act Theory:
1. Illocutionary acts are acts that occur in particular contexts: Illocutionary acts are uses; in different contexts the same locution may be put to different uses.

2. The audience of a speech act (as part of its context) partly determines illocutionary force: An utterance among equals can be different from the same utterance occurring in a hierarchy.

3. The uptake an audience gives can play an important (and sometimes determinative) role in constituting a speech act: Ex: refusals, jokes, bets, questions, etc.

With this in mind, **I show**:

● Through their <u>algorithmic infrastructure</u>, platforms serve as *constitutive intermediaries* that do more than merely connect speakers and audiences, but shape the *total context of interaction* (the 'total speech situation,' as J.L. Austin would call it)

● Through <u>affordances</u>, platforms invite and refuse certain forms of user behaviour, meaning user actions are *co-constituted* by the platform within which they occur.

● Through <u>amplification</u>, platforms' content recommendation algorithms choose audiences to whom to repeat and deliver user-generated content. Platforms should therefore (sometimes) be thought of as co-speakers (of a sort).

---

1. **ALGORITHMS (Platforms as algorithmic intermediaries)**

Tarleton Gillespie, *Custodians of the Internet*:
> "platforms don't just mediate public discourse, they constitute it." … "The moment [social media platforms] did anything other than list users' contributions in reverse chronological order—they moved from delivering content for the person posting it to constituting it for the person accessing it."

The built features of our online environments provide a different context than offline speech, and that **context** *partly constitutes the force of online speech*.

Mary Louise Pratt, "Ideology and Speech-Act Theory":
> "context is not just the backdrop against which a person speaks; rather, the context and the subject mutually determine each other ongoingly. Beliefs, desires, and intentions are seen not as arising out of and attaching to an authentic, monolithic self, but rather as forces that are in play in the situation."

Upshot: Online platforms are therefore **constitutive intermediaries**. Much more than mere conduits, the platforms that host and disseminate our online acts play a co-constitutive role in making our speech what it is.

2. **AFFORDANCES (Features as tools for encouraging desired outcomes)**

> **Affordance**: "the 'multifaceted relational structure' between an object/technology and the use that enables or constrains potential behavioral outcomes in a particular context" (Evans et al 2017).

Or, as **Jenny Davis** puts it "affordances mediate between a technology's features and its outcomes. Technologies don't make people do things but instead, push, pull, enable, and constrain. Affordances are how objects shape action for socially situated subjects."

Seth Lazar: "Algorithmic governance cannot be neutral because it is implicated in what it mediates and because there is no 'natural' baseline to fall back on. So it requires making explicit judgements of right and wrong everywhere"

Upshot: Through various features, platforms shape our speech acts. This occurs in 3 ways:
1. by encouraging and enabling some acts over others;
2. by encouraging and enabling types of uptake that impacts the force; and
3. by themselves being 2nd Personal speech acts, which call for a response (and data).

3. **AMPLIFICATION (Recommendation Systems as features that select audiences)**

Algorithmic ~~amplification~~ recommendation makes platforms co-speakers, alongside the original poster. These processes of curation and amplification take a post and insert it into many new contexts. It is the platform decides which contexts to insert these posts into, and in doing so contributes to the force of the utterance in unique ways.

	Subscription	Network	Algorithm
What a user sees	Posts by those they've subscribed to	Posts by (or shared by) those they've subscribed to	Posts the algorithm predicts the user will like best
Examples	Newspapers, Substack, FB pre-2009, IG pre-2022	Word of mouth, the web, Twitter pre-2016, Mastodon	TikTok, Google Discover, YouTube
What impacts a post's reach	Poster's subscriber count	Both subscriber count and content	The content of the post

[Three stylized models of information propagation from Narayanan (2023)]

When we look closer at Algorithmic Recommendation, we see it is not like a megaphone that amplifies stable speech acts, and instead involves discrete acts of repetition. It is a type of repetition only platforms are capable of. These systems take utterances with indeterminate contexts (locutionary acts) and insert them new, real, contexts (giving them illocutionary force). At the extreme, this may enable online incitement where the context of use is one where the audience gives the utterance uptake that renders that content as a particular speech act.

This is different than how **Jeff Howard** presents the case:
> "My claim is that when a platform amplifies wrongful speech, increasing its visibility, it thereby makes a greater causal contribution to the speaker's wrongdoing—making his principal wrongdoing worse than it would otherwise be. Like the gun vendor who sells the terrorist a larger weapon, enabling him to kill more people, platform amplification enables wrongful speakers to commit a greater wrong."

Recommendation does *not* simply increase the reach of stable speech acts. By choosing the audience, and placing speech acts in new contexts, and by being speech acts themselves, platform recommendations enable something different.

Upshot: With amplification algorithms—or rather, recommender systems—platforms control the audience of an utterance, inserting it into contexts of their choosing, making them co-producers (of a sort) of whatever speech acts that utterance constitutes.

3.1 What Kind of Speech Act is Recommendation?

Javier González de Prado Salas & Ivan Milić:
> "The content of recommendations is not truth-evaluable: when the speaker offers a recommendation, she is not describing things as being a certain way, but rather inviting the audience to do something."

So, Recommendations have **verdictive** and **exercitive** force.
- Recommendations function as **directives**:
  They differ in strength and structure for nearby speech acts (invitations, requests, etc.)

- They are **second-personal speech acts**:
  They are issued by one agent to another: the recommender the recommendee.

- They **presuppose** the recommendee has reason to **trust** the recommender's judgment:
  The recommender is **entitled** to issue their recommendation.

Objection: Don't *extremely thin* recommendations do pose a problem? Yes! But one we can meet by considering where decisions about values are made in the recommendation process.

Arvand Narayan on the (platforms') preference for *implicit* over *explicit* feedback mechanisms:

> "[P]latforms emphasize feedback types that are more frequent. An example of this viewpoint from YouTube researchers in 2016: 'Although explicit feedback mechanisms exist on YouTube (thumbs up/down, in-product surveys, etc.) we use the implicit feedback of watches to train the model, where a user completing a video is a positive example. This choice is based on the orders of magnitude more implicit user history available.' This is generally true across platforms, and over time, there has been a shift to "implicit" forms of feedback where the user action is minimal.

If (as Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi) put it, the problem lies in "the **opacity** about **which and whose values** are at stake in recommender systems," does this mean that greater transparency and user-agency can relieve some of these issues? Maybe!

Upshot: Recommendations get the speaker on the (moral) hook.

5. CONCLUSION

The Upshot (of all this):

> Social Media Platforms are *not* mere conduits; they are **constitutive intermediaries** of our online speech acts. They shape our speech acts through their choices regarding context, audience, and uptake.

> They speak themselves through their **recommendations** and various **affordances**, both of which **express their values** about how users should use the platform.

> They share **responsibility** for what occurs on their platforms, not because they are (merely) complicit, but because they are **co-producers** of the acts that occur there.

> They represent **something new**. We must collectively come to a decision about what that thing is, and how we want it to act.

In sum, **Speech act theory** provides a useful set of analytical tools that help to uncover aspects of online speech. It opens up lines of analysis that enable a greater understanding of what actions are at our disposal in the many contexts of online speech, and, more fundamentally, reveals the conditions of the overall environment that enable those actions.

This matters for online speech because, more so than many other mediums, that 'total speech situation' is one that is manufactured and controlled by an intermediary. That position generates a lot of power (what Lazar calls 'intermediary power') and speech act theory illuminates this more fully, showing how this power is made manifest.

Comments welcome! email: michael.barnes@anu.edu.au