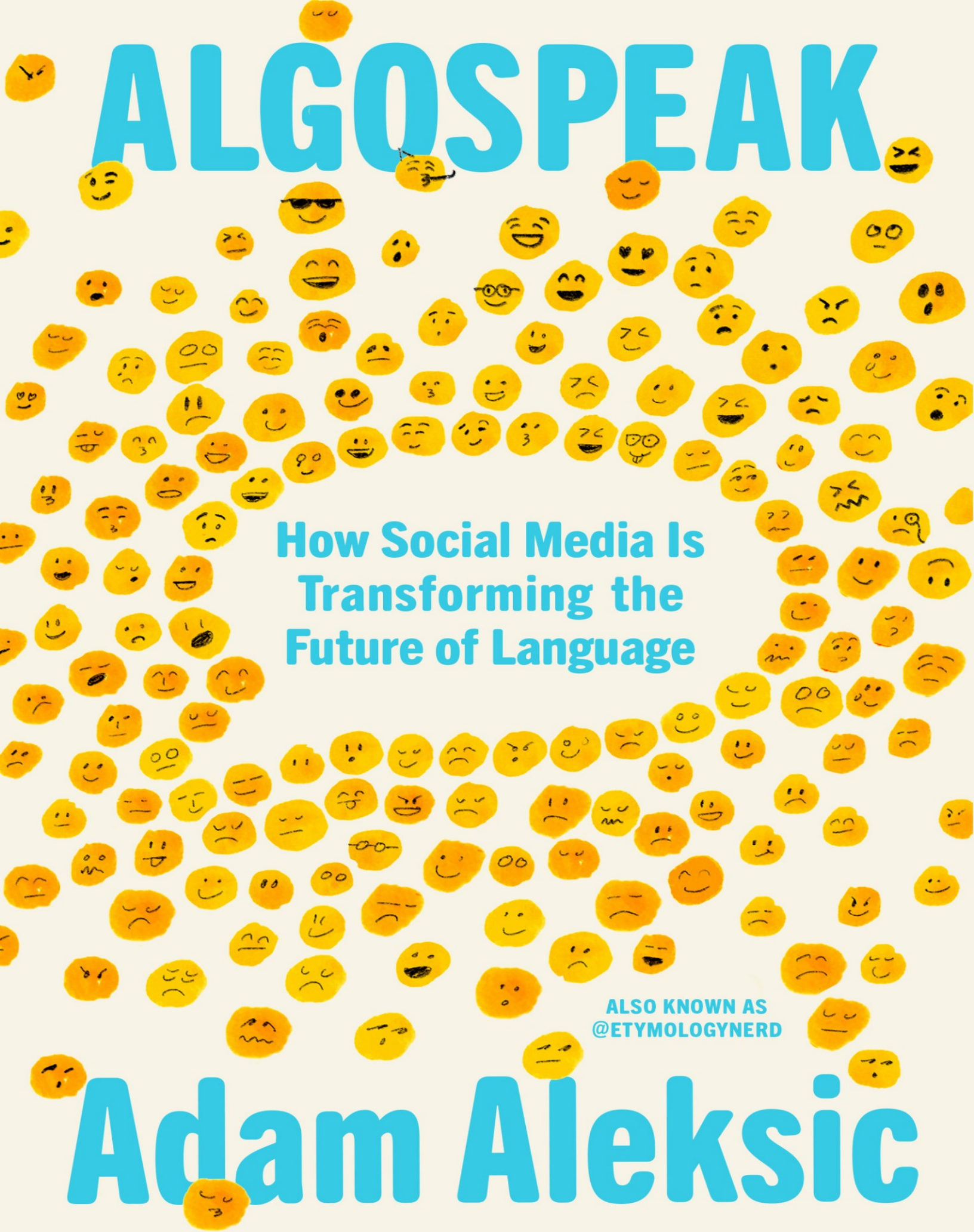
ALGOSPEAK



How Social Media Is Transforming the Future of Language

ALSO KNOWN AS
@ETYMOLOGYNERD

Adam Aleksic

ALGOSPEAK

How Social Media Is Transforming the Future of Language

Adam Aleksic

ALFRED A. KNOPF • NEW YORK • 2025

A BORZOI BOOK

FIRST EDITION PUBLISHED BY ALFRED A. KNOPF 2025

Published by Alfred A. Knopf, a division of Penguin Random House LLC, 1745 Broadway, New York, NY 10019.

Library of Congress Cataloging-in-Publication Data

Names: Aleksic, Adam, author

Title: Algospeak : how social media is transforming the future of language / Adam Aleksic.

Description: New York : Alfred A. Knopf, 2025. | Includes index.

Identifiers: LCCN 2024051333 | ISBN 9780593804070 (hardcover) | ISBN 9780593804087 (ebook) | ISBN 9781524712907 (open market)

Subjects: LCSH: Language and the Internet | Internet users—Language | Youth—Language | Social media

Classification: LCC P120.I6 A54 2025 | DDC 402.85/4678—dc23/eng/20250411

LC record available at https://lccn.loc.gov/2024051333

Ebook ISBN 9780593804087

Cover illustration by Ben Hickey

Cover design by Tyler Comrie

penguinrandomhouse.com | aaknopf.com

The authorized representative in the EU for product safety and compliance is Penguin Random House Ireland, Morrison Chambers, 32 Nassau Street, Dublin D02 YH68, Ireland, https://eu-contact.penguin.ie.

ep_prh_7.3a_152340698_c0_r0

CONTENTS

To Kirkland

INTRODUCTION

.....

Why Your Kids Are Saying "Unalive"

You're a die-hard Nirvana fan, and it's the thirtieth anniversary of Kurt Cobain's suicide. You want to pay your respects, and you heard Seattle's Museum of Pop Culture is putting on an exhibit to commemorate the lead singer. When you get there, however, you're shocked by the first placard under the display. It doesn't say that your favorite singer killed himself. It doesn't say that he "committed suicide." Instead, it tells you that "Kurt Cobain un-alived himself at 27." Huh.

It's understandable if you would've felt upset in this situation. In fact, so many museum visitors were riled up about it that curators were forced to replace the placard a few days after the faux pas went viral. The term "unalive" seemed to trivialize the topic of suicide, making light of a death still felt by many. Even more jarringly, it represented the first notable offline use of a word made for dodging internet censorship.

"Unalive" might have been surprising to casual museumgoers, but by that point it was already quite normal to middle school teachers. That's because, if you spend enough time around seventh and eighth graders, you're virtually guaranteed to hear them use the word as a synonym for "kill" or "commit suicide."

It might be a bully in a hallway telling someone to "just unalive yourself." It might be an edgy teenager saying they'd rather unalive themselves than clean their room. Or it might even appear in more formal contexts, such as student essays on Hamlet's contemplation of unaliving himself or classroom discussions on the unaliving that happens in *Dr. Jekyll and Mr. Hyde*.

These aren't hypothetical situations. Each of those examples was indeed overheard by one of the more than seventeen hundred middle school parents and teachers that I surveyed while researching this book. It might sound ridiculous to older audiences, but "unalive" is very much alive. In fact, it's the future of the English language.

For a relatively recent word, "unalive" can be applied to an impressive range of scenarios. It can show up in informal speech, such as the bullying example, or formal speech, such as the Hamlet example. It can occur in ironic or meme contexts: One teacher overheard their student say, "Let's go unalive some sandwiches."

Most important, the main function of "unalive" appears to be euphemistic. Many adolescents use it when they're uncomfortable talking about the concepts of death and suicide, since "unalive" sounds like a less scary word. While this development has caused much hand-wringing among adults concerned about trivializing suicide, it could very well be positive for mental health education: One 2023 Associated Press article argued that "unalive" makes it easier for counselors and students to have "meaningful discussion" over difficult topics,[1] and many of the teachers I talked to echoed this idea.

In some cases, students are even learning the word "unalive" before the word "suicide." It's what their classmates are more comfortable saying, and since it gives kids a new way to express themselves, "unalive" will probably only continue to grow in popularity. It's already begun creeping into more formal, educational settings, such as the Hamlet essays or academic discussions on Dr. Jekyll unaliving himself, because students genuinely perceive it as a classroom-safe alternative to its heavier synonyms. In a side placard, the Museum of Pop Culture gave this rationale for using the word:

> "Unalive" has sparked constructive conversations, especially among young people, addressing issues such as depression, anxiety, and suicide.

In many ways, this is nothing new. We have been euphemizing death as long as we've had language. The verb "to decease," for example, comes from *decessus,* the Latin word for "departure," which was introduced as a more agreeable alternative to the existing word for death, *mors*. Similarly, the verb "to die" was originally an Old Norse euphemism for the Old English verb *sweltan,* which was, in turn, likely also a

euphemism for the even older English word *diegan* (related to “die” but then forgotten long enough to be reintroduced in a new form).

Clearly, even the stoic Romans and fearless Vikings got as queasy about death as modern middle-schoolers. It’s only because we’ve been constantly running linguistic circles around our mortality that we still use phrases like “passed away,” “left this world,” and “departed.” In that regard, “unalive” is just another step in this ever-evolving process.

The crucial difference, however, between “unalive” and “decease”—and the reason it came across as so offensive—is the context in which “unalive” emerged. The word first appeared in its modern verb form in a 2013 episode of *Ultimate Spider-Man* and then circulated on a few meme sites without ever gaining much offline traction.[2] After that, it might have been relegated to the ever-expanding void of obscure internet history had it not been for the Chinese government.

Starting in 2019, China began enforcing its censorship policies such that social media companies would be held accountable for short-form video content generated by their users. One of those companies, ByteDance, had a popular app called Douyin that was subject to these regulations; to adhere to these new censorship policies, it began aggressively blocking any content critical of China, including discussion of topics such as the Hong Kong protests and the Xinjiang internment camps of the Uyghur people. This was all accomplished through a new “sensitive words” tool that drew on an extensive library of high-risk keywords to identify and remove any instances of those keywords being spoken or written or somehow appearing on-screen. The keywords tool was then combined with a string of other AI-driven processes to create a personalized stream of recommended content we collectively call the algorithm.[*1]

This “algorithm” was then applied to Douyin’s international counterpart, TikTok. The exact vocabulary lists were most likely tweaked for the new markets, although we can’t know for sure, because everything about them is shrouded in corporate secrecy. There is substantial reason to believe that ByteDance has suppressed anti-Beijing sentiment on TikTok,[3] and that it has previously censored content related to

LGBTQ+ themes and other sensitive topics such as politics, pornography, and self-harm.

An unfortunate result of this policy is that it became very difficult for online creators to have earnest conversations about topics that might incidentally register as violations of the TikTok community guidelines, such as suicide prevention. Knowing that their video might be removed for talking about their own experiences, some influencers opted to use lesser-known alternative words. This is how "unalive" made its way from a *Spider-Man* meme to become a widespread synonym for "kill."

Popularized through a few viral videos in early 2021, "unalive" exploded in usage and became a widespread online term from 2022 onward.[4] Eventually, it made the leap from TikTok to other contexts, such as video game streaming, where apps like Twitch also restrict discussion of suicide. At a certain point, some people started saying it primarily out of habit, or because they saw other people saying it. That's how "unalive" was able to make a second leap, from online to offline usage.

The middle-schoolers obviously don't know this. They aren't saying "unalive" to avoid algorithmic censorship. They do so because they see other people using it, and it genuinely does serve a very useful euphemistic function. Many might not even be aware of its origins: A majority of the teachers I surveyed were sure that at least some of their students didn't know where "unalive" came from. They simply hear it from content creators, or from their peers who are more active on social media, and then start using it because it's a helpful, versatile word that can be applied in a variety of situations. This kind of positive feedback loop ensures that, with time, fewer and fewer people will know the story behind "unalive." Eventually, more institutions like the Seattle museum will pick up the word to keep up with the youth, until it eventually becomes so normalized that its etymology is completely forgotten.

Honestly, fair enough. You probably didn't know where the word "decease" came from, unless you studied Latin or are some kind of etymology nerd. But I think it's safe to say that "decease" didn't happen just because it was impossible to carve the word *mors* into an ancient Roman tablet. We're entering an entirely new era of etymology, driven by the invisible forces behind social media and its algorithms.

.....

Traditionally, this kind of evasive speech is called *algospeak*. Fittingly, the word seems to have been created on social media: The earliest references to the term for "algorithms shaping how you speak" are from scattered tweets and TikToks before Taylor Lorenz popularized the concept in a 2022 *Washington Post* article.

The word was sorely needed, because algospeak is everywhere. A 2022 TELUS International survey found that three-quarters of Americans have encountered some version of algospeak online, with 30 percent actively using coded language to circumvent content moderation. I would venture that the numbers are a lot higher, since many people unknowingly come across or reproduce these words in offline interactions. Especially now that we have millions of middle-schoolers using algospeak in their daily lives, we'll only see it continue to grow more relevant.

As a linguist, content creator, and human being, I've always had a very strange relationship with language on social media. It's my research interest, my livelihood, and my social life, and I perceive it to an almost problematic degree. I can't record a video without fixating on my own word choice, and I can't enjoy an Instagram Reel without also linguistically analyzing some aspect of it. I find myself constantly noticing how the format of social media affects my expressive decisions and observing where they've probably impacted those of other creators.

Well beyond censorship avoidance, I see everybody mold their speech around algorithms anytime language spreads on social media. The personalized recommendation structure of the TikTok "For You" page was a critical vehicle in helping "unalive" expand beyond the context of a niche *Spider-Man* reference. Since the platform used keywords to identify people passionate about mental health, it was able to bring those people together to form an online community—one that needed a way to spread resources and talk about their shared experiences without worrying about their posts being removed.

Once the mental health community repurposed "unalive" to suit their needs, the term slowly filtered into other communities on TikTok, since the boundaries of in-groups are much more porous on algorithmically based applications. As more people encountered the word and found it practical, the concept exponentially spread until it fully became a part of TikTok's culture. To be on the platform was to be familiar with

its algospeak. Eventually, it spread to other social media, and then over to offline settings, or IRL.

"Unalive" also shows how the diffusion of words has become indistinguishable from the diffusion of memes. In its earliest days, the concept spread only through joke contexts until it found a place as a serious euphemism. Yet this is now how all concepts spread on social media: We find something compelling, we share it, and it goes viral. This usually happens in the form of memes, which the algorithm picks up on and helps to circulate. Social media platforms want to promote the most compelling content possible, so it makes sense that words will reach us through maximally compelling mediums like memes rather than something more serious. "Unalive" is far more likely to spread today than boring, traditional alternatives such as "passed away."

To me, the old definition of "algospeak" is far too limited in scope. Algorithms aren't just giving us euphemistic substitutions; they're playing a *defining role* in every aspect of modern etymology. They shape who gets exposed to certain words, how those words spread, and how popular the words eventually become. Language and memes and metadata are one and the same—all of it shaping our vocabulary and identities.

On some level, we can tell this is happening. The reactionary backlash to the Kurt Cobain "unalive" exhibit placard wasn't just about censorship. It demonstrated our wariness over our own algorithmic dependence, throwing it into sharp relief through an elevated museum setting that was supposed to be separate from all that. Other euphemisms would not have been nearly as controversial, but this was an *online word* being used *offline*. Clearly, that boundary has started getting blurred.

.....

About half a dozen times throughout human history, we've reached inflection points that fundamentally changed the nature of our communication. The shift to writing allowed language to evolve beyond the purely spoken medium, marking the division between prehistory and history. The printing press increased literacy rates and brought about the standardization of language. The internet itself marked another inflection

point, broadening public information and allowing the written replication of informal speech.[*2]

We're now at a new inflection point, characterized by personally recommended, short-form video content. All channels of communication—advertising, education, news, entertainment—are gravitating toward the medium by necessity, because that's what people have been conditioned to consume. If you want to go viral, you must make short-form video, and that means that you must mold your language to the constraints of short-form video platforms.

In many respects, the algospeak era gives us a new medium to retell very old stories. It's incredible that we're seeing time-honored patterns—like the continued ways to euphemize death—play out in a novel context. It's still the same process, just happening in a new setting and happening much faster than ever before. While it might have historically taken words like "decease" hundreds of years to catch on, "unalive" did the same thing in just a decade, thanks to the new medium of TikTok and its subsequent copycats, like Instagram Reels and YouTube Shorts.

We're not just getting new words, either: The new structure of social media has given rise to entirely new accents, dialects, grammatical rules, and morphological processes. It's driven by community guidelines, but also by the nature of virality, the tricks influencers use to grab your attention, and the way your "recommended" feed is formatted.

Shockingly, for something with such an enormous impact on our day-to-day lives and language use, there's very little research or literature covering the algorithmic inflection point. Part of the problem is that it's so new. Researchers might find it hard to keep up with the fast pace of language change, especially if they're of an older generation. By the time these changes reach widespread use among younger demographics, adults are only barely starting to catch on. Even when they do, they might gripe about how today's youth are being corrupted, without regarding their words as serious linguistic developments. I believe it will be decades down the line, when words like "unalive" are completely normalized, that they'll ever be rigorously studied by traditional academia.

Another problem is the sheer volume of information to process. With more than seven billion internet users uploading over 400 million terabytes of media every day (the equivalent of 20,000 Libraries of Congress), we're generating an obscene amount of data for linguists to sift through. This can easily be overwhelming. No matter how much work we put in, we're never going to come close to capturing the infinitude of linguistic variation that's out there in plain sight. Even for the few hundred slang words that seem most culturally significant at this moment, it'll take decades for etymologists to fully catch up. Online user-generated repositories like Urban Dictionary, Know Your Meme, and Wiktionary are genuinely some of the most useful records we have of how language is evolving, but you never really know what you can trust.

Last, there's the transient nature of pop culture. It can be very hard to tell which meme is here to stay and which meme is a fad; this could discourage researchers from investing time into something that might not be relevant for long. As I write this in the fall of 2024, much of the slang I mention is already outdated. What seems current to me now might be passé by the time you read this book. If you reread this down the line, you'll notice the same thing. Some words will be gone, while new words will have emerged that you could never have predicted or imagined. And that's normal.

One thing is constant, however: this overarching pattern of algorithmic media causing linguistic innovation. Personalized recommendations aren't going anywhere for the foreseeable future, because the social media landscape is too competitive. Any country that has banned TikTok, for example, immediately saw rival apps like Instagram swoop in to dominate user attention. Short-form video is simply the most addictive medium we have, which means "the algorithm" is here to stay. This is why I think it's absolutely worth talking about even the most fleeting words. We must, if we want to understand where we're headed as a society.

"Unalive" is the tip of a colossal iceberg. Beneath the icy water, there is a hulking mass of creative techniques, trends, memes, identities, echo chambers, and new ways to commoditize language that only could have existed in the new algorithmic context. It's about time we started talking about it, so I invite you to join me in navigating our new linguistic landscape. Throughout this book, we'll dive into every corner of social media and together uncover the beautiful, chaotic idiosyncrasies of modern language. We'll find out who we are, who we're becoming, and what that means.

Great. Let's play Whac-A-Mole.

Skip Notes

*1 I'll keep calling this collection of processes *the algorithm,* since that's how we talk about it colloquially, but remember that it's a little more complicated under the hood.

*2 David Crystal and Gretchen McCulloch both covered this inflection point phenomenally in their books *Language and the Internet* and *Because Internet,* respectively.

1

How to Play Linguistic Whac-A-Mole

Remember whac-a-mole? The slightly unhinged arcade game where new characters keep popping up no matter how many times you smack them with your mallet? It might seem ridiculous to use coin-operated violence as an analogy for the serious linguistic changes we're experiencing online, but Whac-A-Mole truly is the perfect metaphor for explaining how humans react to censorship. As soon as a word is banned, we find a way around it; that is, until content moderators catch wind of the new word and ban that, too. Then another word pops up, and the cycle repeats, trapping users and platforms in a never-ending loop of new spellings and substitutions that disappear once the algorithm catches on and the mallet comes down again. The faster and better the moderation tools are, the more words will be created. This is because the underlying idea—and our desire to talk about it—remains.

People have been playing this game since the early days of the internet. Frustrated with text filters on bulletin boards in the 1980s, netizens turned to "leetspeak," a hacker dialect characterized by creatively respelled words. Since terms like "suicide" were censored in some chat rooms, leetspeakers wrote out coded replacements like "5U1C1D3" in much the same way that people began using "unalive." If the chat-room moderators caught on, they could just change the spelling to another easily recognizable form, like "$U!C!D€."

The tools being used have since grown more sophisticated, but the underlying process remains the same. Rudimentary word identification scripts might have been upgraded to fancy, AI-powered algorithms, but censorship is still driving linguistic creativity. If anything, the algorithmic era has spurred more innovation, because the game is happening faster, with more guesswork involved. Due to intellectual property

concerns, not much is known about how these apps actually recommend content. Influencers are often subject to the whims of unknown and unfair criteria, with only opaque community guidelines as reference points.

This marks a major difference between leetspeak and algospeak: You could always immediately tell when leetspeak worked, because you would see your comment successfully posted onto a message board. For algospeak, however, the goal is to make it onto a user's "recommended page," and it's much harder to tell when you've successfully done that. Videos with sensitive keywords aren't always removed outright. Sometimes, they'll be "suppressed," or shown to fewer followers. Creators can also be "shadowbanned" without warning or notification. We—and I'm including myself here—receive very little communication from the platforms, so we're not sure whether videos do poorly because they're bad or because they're being censored. Understandably, then, we'll err on the side of caution when it comes to euphemization.

Creators *can* sort of tell what does and doesn't get onto the recommended page by looking at our video analytics. If I usually get most of my views from the recommended page, and then all of a sudden a video is getting views only from the "followers" feed, that's a semi-reliable indication that some part of my video made the algorithm unhappy. Through this kind of trial and error, influencers can extrapolate a pretty good idea of what the algorithm rewards and penalizes, and it's in this context that "unalive" was forged. When the word "suicide" wasn't getting views, people turned to the next best term to tell their story.

While linguistic innovation like this is an exciting and normal thing, it's reasonable to be alarmed at the way community guidelines are shaping important conversations, especially as people increasingly turn to short-form video for news or advice. Oftentimes, the mysterious rules governing social media are arbitrary or outright discriminatory. TikTok has historically been proven,[1] for example, to artificially suppress videos by "ugly," old, and poor creators, because they're not as appealing to new users. This means that it's often difficult to include larger audiences in discussions about things like disability, age, and income inequality. Nevertheless, we have no choice but to play the game and tiptoe around community guidelines wherever we can.

.....

In 2022, the Charles Dickens Museum began a desperate social media campaign to get itself un-shadowbanned from TikTok. Whenever users would search for the museum's account, nothing would show up. Instead, they would be cautioned[2] that "this phrase may be associated with behavior or content that violates our guidelines" and that "promoting a safe and positive experience is TikTok's top priority."

Of course, the Charles Dickens Museum wasn't doing anything wrong. They were mostly posting house tours or excerpts of old letters. Instead, the problem was with the TikTok algorithm, which was flagging the museum's videos as obscene because they included the keyword "dick." For all its fancy high-tech machine learning, the algorithm had fallen victim to a classic internet pitfall: the Scunthorpe problem, named for an English village where residents discovered they were unable to create AOL accounts because their hometown contained the word "cunt."

Following an intense #FreeDickens campaign on Twitter, TikTok eventually agreed to unblock Charles Dickens–related search terms. However, it still remains difficult to curse—intentionally or accidentally—on any platform. While your videos won't get removed outright, they'll often be suppressed in search just like the museum's. Especially if you're cursing too much or too severely, your content will be hampered by the algorithm. If a video is eligible for the recommended page but still contains mature language, TikTok and Instagram will prevent it from appearing in clusters of similar videos in a user's feed, which means that it'll be pushed to a smaller audience than its work-safe competitors.[3]

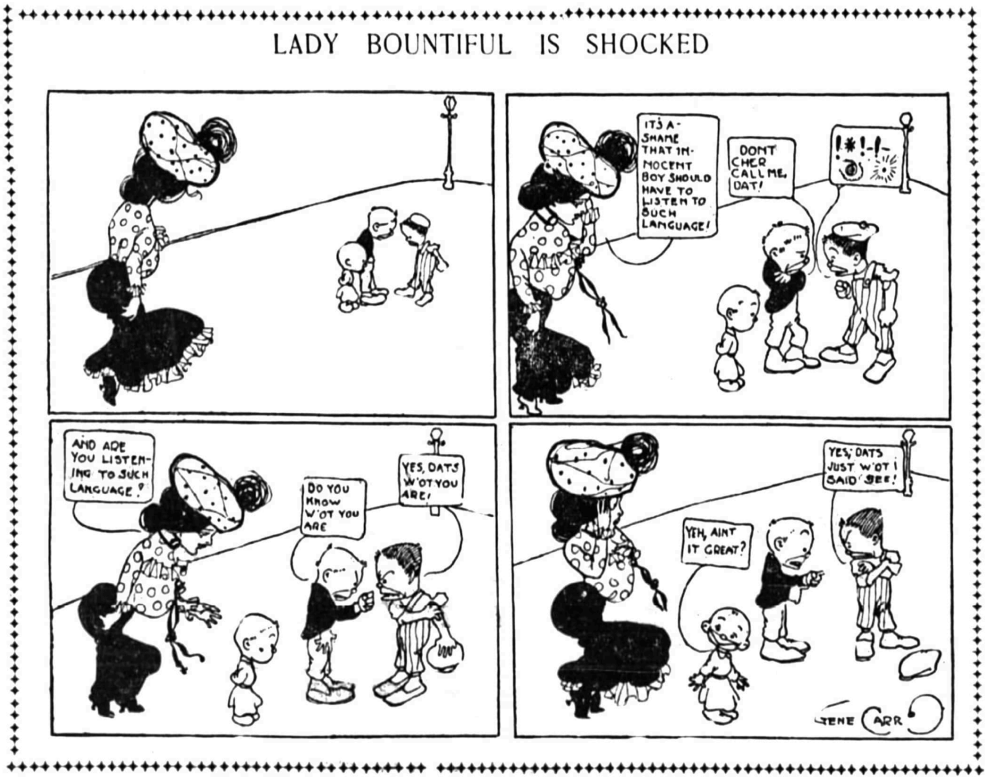
Likewise, on YouTube, creators posting videos with severe or repeated profanity have to contend with "demonetization," where they lose the ability to earn revenue from advertisements on their content. This is especially scary to influencers whose livelihoods depend on a steady income stream from their videos.

All this means that we have a lot of reasons to respell our swear words creatively online. If you look up the keyword "bitch" through TikTok's search function, for example, you'll likely encounter variations such as "btch," "b!tch," and "b*tch" in the video captions. Same with "fuck," which will probably give you "fck," "fvck," and "f*ck"—none too different from the leetspeak letter substitutions of the past.

The practice of respelling offensive words is a centuries-old tradition known as *bowdlerization,* named for the Englishman Thomas Bowdler, who is mainly

remembered for publishing some egregiously family-safe edits of William Shakespeare's plays.

Self-bowdlerizing to avoid media constraints is not new: People have been doing it since at least the days of early newspaper comics, where sequences of graphic characters called *grawlixes* are still used instead of swear words to circumvent stringent syndication standards. The earliest known example is from this 1901 *Lady Bountiful* comic by Gene Carr:



Over time, some grawlixes got less thinly veiled as cultural norms against profanity loosened up. Cartoonists began drawing on symbols with a visual similarity to the letters they were replacing: "@##" and "$#*!" are now industry standards for the words "ass" and "shit." Fast-forward to today, and influencers are re-creating this process in social media captions. Words like "fvck," "b!tch," and "@ss" are born out of the same motivations, drawing on the same cultural tradition of bowdlerization.

Until the twentieth century, the preferred method of bowdlerization across all media formats was the double em dash (——), typically replacing the entirety of a