

The Audio Revolution, Part Two

Alex Danco

17-22 minutes

Two Truths and a Take

Every Sunday, from Alex Danco



Today's newsletter is Part Two of a two-part series; Part One came out last week. If you haven't read it yet, [go read it here first](#).

When we said last week that hot media *create space* for hot messages, or create a kind of stage on which they succeed, what do we mean by that? Where is that stage? Well, we are the stage. Our attention and comprehension is where information “happens”.

So in order to understand what headphones and audio are doing to us, we need to take a closer look at us: that is to say, our brains.

Information and the Brain

Brains face an engineering challenge: how to deal with sensory input streaming in, in real time. It's a speed problem. Individual neurons in your brain can each take somewhere in the tens or sometimes hundreds of milliseconds to integrate and transmit signals between each other. Even basic neural circuits can comprise dozens of neurons. Without some way to speed this up, complex sensory integration or motor output would be impossibly laggy.

We use something called [feed-forward](#) processing in order to speed things up. Feed-forward processing is useful when you're interpreting inbound information that's familiar or predictable. If you're reading the sentence: *Somebody once told me the world is gonna roll me, I ain't the sherpest tool in the shed; she was looking kind of ocean* (Wait, what?)

Feed-forward mode:

High Resolution mode:

Low engagement, high participation

High engagement, low participation

Cool

Hot

Visual and/or multiple senses,
discussion format = feed-forward,
gap-filling, cool sensory processing

Uni-modal sensory input, auditory
& language, single shot input
= high-resolution processing

What happened? You began the sentence, and then your brain picked up on a pattern it recognized: Smash Mouth lyrics. Then your reading sped up – you already know those lyrics, so you *fed them forward* into your sensory processing stream. You start skimming: reading in low resolution and filling in the gaps. But then, you hit the word *ocean* and slammed to a stop: it didn't fit the model you fed forward. Better go back to reading one word at a time.

Once we start following the known All Star sequence, each additional word contributes almost zero new information, because they resolve no uncertainty. (Filling in *The World is Gonna* _____ with Roll Me happens automatically). But the word Ocean was new information. You sense, “There’s uncertainty to resolve here”, and flip back into high-resolution information processing, which is much more discriminatory.

Feed-forward prediction is one of our brain’s critical information processing tools. We rely on it continuously, at every abstraction level, from basic raw input up to executive function – particularly for our eyes. Our default way of processing the world isn’t taking it all in finely discriminatory hi-fi, it’s continually assembling and filling in our understanding of the world with what we expect is there.

In real sensory perception, you’re continually making subconscious, probabilistic judgement calls about when to switch into intense, high-res inspection versus when to keep scanning and gap-filling in low resolution. If you look back at those lyrics, you’ll see that I actually wrote “Sherpest” instead of Sharpest, but you may not have picked up on it. It’s a small error, so it may not have flipped the switch. Or maybe it did! With neuroscience, everything is just a probability.

Hot and Cool Brain Muscles

The reason we’ve gone through this little neuroscience lesson is to build towards an important point. The brain continuously triages inbound sensory input into our low-resolution, fed-forward, gap-filling stream and into our high-resolution, information-saturated stream. This should ring some bells: sure sounds like Cool and Hot. And it is.



One of the major differences between hot and cool media that we only briefly touched on before is McLuhan’s classification of hot media saturating one single sense, whereas cool media often integrates multiple senses, filling in a picture from many inputs. How come?

The neurological explanation is illuminating. Sensory input is processed in two different ways: uni-modally (vision only; audio only) and multi-modally (integrating multiple senses together into a complete picture). We don’t totally understand why, but we believe that our uni-modal sensory processing pathways are **more sensitive to uncertainty and “New Information”** than our multi-modal pathways are. Our neural circuitry dedicated to integrating multi-modal sensory information is less willing to throw the switch into high-resolution, finely discriminating information processing. It prefers to scan in low-resolution and fill gaps. It’s cooler.

Meanwhile, not all senses are created equal. Inbound audio, particularly human speech, is particularly sensitive at triggering the “There’s information to resolve here” mode of sensory processing. Written text, which passes through our language areas (evolutionary speaking, an audio domain) is pretty sensitive too. There’s also a difference between discussion versus monolog formats: cool dialog, where information is communicated in gaps and pauses, asks for more participation (feed-forward gap-filling) than a single-shot, high resolution stream of inbound information.

Now we’re ready to understand the impact of Hot and Cool at ground-truth level:

Cool sensory perception and cool media are **low in engagement but high in participation**. We are operating in gap-filling mode: doing relatively little *engagement* with the media (we’re only pulling in a low-resolution sample) but a lot of *participation* with the media (we’re actively filling in the gaps ourselves, and operating in feed-forward mode).

Hot sensory perception and hot media are **high in engagement but low in participation**. We’ve switched out of feed-forward scanning mode: doing a lot of *engagement* with the media (we’re intensely processing a high-resolution inbound sensory stream) but not much *participation* (because there are no gaps to fill in).

Feed-forward mode:

High Resolution mode:

Low engagement, high participation

High engagement, low participation

Cool

Hot

Visual and/or multiple senses,
discussion format = feed-forward,
gap-filling, cool sensory processing

Uni-modal sensory input, auditory
& language, single shot input
= high-resolution processing

If you remember one thing from this essay, remember this: hot sensory processing and cool sensory processing are like muscles. The more you use them, the stronger they get, and the stronger they get, the more we use them. We used to think that our neural circuits were relatively fixed by adulthood, but we now know better: they're highly adaptive, and they strengthen and synchronize with repeated use.

As you use cool neural circuits, you create a cool stage that will easily and fluently accommodate cool media and cool messages. As you use hot neural circuits, you create a hot stage that intensely and eagerly accommodates hot media and hot messages. The stage is you.

So how about those headphones?

Old and New Radio

Radio is a perfectly hot form of media. It maxes out all three dimensions we care about: it saturates one single sense, it's spoken audio that's high in information density, and it's a uni-directional blast of information. Everything about the radio medium pushes our brains towards high-resolution, high engagement, low participation mode.

American radio has an interesting history. The earliest days of amateur ham radio gave way to Radio as Big Business, led by the Radio Corporation of America in the first half of the 20th century. (Tim Wu's book *The Master Switch* is a good intro in context with other media.) But as television took over much of centralized broadcasting, radio reorganized itself into a vibrant local kaleidoscope of programming: music, weather and traffic reports and especially talk radio.

From liberal pockets on the coast, it's easy to miss how popular and influential talk radio is in America. The average American adult reportedly [listens to an hour and a half of radio a day \(!\)](#), of which a heavily skewed 15% or so is talk radio. It's more politically varied than you think, especially if we include internet radio and podcasts, but AM radio has been the power base of the American Right for a long time. It's an intimate, private format: the host is speaking directly to you, in high definition. The most important place where radio is listened to isn't in public or even at home, it's in your car: a private environment where you and Sean Hannity battle traffic together.

Headphones recreate that environment: a completely private space, just for the two of you. Mobile phones and the internet put anybody in the world in your pocket, and then headphones complete the gap. Remarkably, the rise of streaming audio (music, podcasts, audio books, internet radio) has only slightly dented radio listening statistics in America: almost all of this new streaming is *additive* to the audio we were already listening to.

Next time you're out, look around at how many people are inside their headphones. All of that audio, all day long, is layered on top of [a world of escalating loudness](#). All of this audio stimulation is doing something to us. Even benign background music has an impact. It's repeatedly juicing our Hot sensory processing brain circuits, shifting that probabilistic balance a tiny bit, away from cooler, participatory, convention-following, feed-forward processing and towards something closer to an alarm state.

Any individual hour of audio has a negligible effect. But hundreds of them? They add up. That's why cool media environments make us receptive to cool messages, and hot media environment makes us receptive to hotter messages. With the internet, we find it. The most important modern audio institution is the new elephant in the room, and it's giving us what we want. It isn't internet radio, nor is it podcasts. Most people don't even realize that they're an audio company.

It's YouTube.

The scale of YouTube is ridiculous. YouTube [reports](#) 1.9 billion monthly logged users, with over a billion hours of content consumed daily. The majority is on mobile. It is the world's second largest search engine and second most visited website, after Google. 400 hours of content are uploaded to YouTube every minute. Nothing else is this big, except Facebook.

One of the biggest misconceptions of YouTube is thinking of it strictly as a video or visual product. Yes, it's a video player; and yes, it's true that there are some specific verticals within YouTube, like gaming or beauty, that are image-heavy. But YouTube's real heritage and impact aren't visual. YouTube is amateur radio, at mega scale.

YouTube built an frictionless platform for people to broadcast and communicate with one another. But it didn't necessarily make content creation easier. It still takes effort, expertise and likely a budget to create content that visually communicates meaningful information. But it's trivially easy to create content that communicates audibly. Just hit record, point the camera towards you, and start talking. Most of the signal coming out of YouTube is people talking.

I really wonder what percentage of YouTube is consumed audio-only, or at least audio-first (the visuals are playing, but the viewer isn't really paying attention to them). I bet you it's way higher than people think. The ability to keep YouTube playing while switching apps on mobile is a major selling point for their premium service; that's explicit YouTube-as-radio use. We know that Music on YouTube is huge, but it's not what I'm talking about. I mean: what percentage of all YouTube content, and of all streaming time, is content that's mainly someone talking, saturating one single sense – your ears – and not much else important is really going on?

If 10% of YouTube consumption falls into this category (and I'll bet you it's higher!) that's 100 million hours of New Radio consumed every day. And this is not lukewarm stuff like Instagram or even loud, indignant rants or conspiracy hoaxes shared on Facebook. This is radio: the format for communicating what you *really* mean.

As I write this, at a table in the library, someone just in front of me has a YouTube tab open in the background, streaming Ben Shapiro. Courteously for the rest of us, he's wearing headphones. Their conversation isn't for us. It's private.

Headphones and America

People accuse Facebook and Twitter of being misinformation platforms that sway elections, but I don't really buy it. I think they're mostly lagging indicators for how people already feel. The kind of urgency that really changes minds isn't a feeling we learn with our eyes. We learn by hearing it: in intonation, in phrasing, in private, in our car radios and headphones. (Our language use betrays this: we use the word "See" to mean "To check out", whereas we use the word "Listen" synonymously with "to pay close attention.") The more we use our ears, the more quickly we're going to pick up on that urgency. The medium IS the message.

When every American put headphones on their ears and then connected those headphones to the internet, should we be surprised that our national politics, values, and discourse are shifting away from cooler, open values and towards hotter, closed ones? I don't just mean the MAGA movement, by the way. Bernie Sanders and Elizabeth Warren are also hot candidates with hot messages, and it's their moment. (Warren is the first candidate I can remember who uses lengthy, block text writing – also a hot medium – as a legitimate political communication tool.)

There's a particular kind of message that hot media consumption primes really well, and especially resonates over audio, that has come to dominate the political conversation in America. The message is *discrimination*. I mean that not just in the way we usually use the word discrimination (as in, exclusionary prejudice) but in a broader sense. In our lifetimes, modern liberalism has built up a cool, open society whose overarching aspirational value is equivalence: the idea of the "level playing field", both economically and socially. It's given us equal rights and anti-discrimination laws, on the one hand, and free trade on the other.

The current backlash against liberalism, which is especially well-articulated in the MAGA crowd (where Trump is a perfect orator), is a backlash against this enshrinement of equivalence and supposedly level playing fields – both economically and socially. Make America Great Again really means *Let America Discriminate Again*. "You're telling me we're supposed to believe there's no difference between X and Y? You and I both know there is *clearly a difference*." You can fill in X and Y with "Citizens versus non-citizens." Or, if you prefer, with "Regular people and rich people."

This message lands harder over audio than it does in text, video or any other format. On the left or on the right, it doesn't matter: the specific message varies, but the mechanism is the same. Headphones create a private space for that finely discriminatory message, and a channel through which it resonates the loudest. It shouldn't surprise anyone that audio has become the most powerful new format for political action on both the left and right in America today: on the left, podcasts; on the right, YouTube.

In *Understanding Media*, Marshall McLuhan spends time ruminating on what happens when hot and cool media enter societies for the first time, or in new ways. The printing press, which quickly spread hot, printed text through Europe, was a pretty big factor in the subsequent centuries of continuous war that followed. More recently, radio has done the same.

Interestingly, McLuhan speculates that England and America were spared from the traumatic effects of hot radio (in contrast to Weimar Germany) because we'd already been "vaccinated" by our higher literacy rates and prevalence of print media. I'm quite sure that the biggest consequences of the Audio Revolution won't be in the United States. They'll be in the developing world. I can't speculate at all on what they'll be, as that's truly beyond me, but a world without headphones almost certainly turns out differently.

Less Audio

A few months ago, I stopped listening to podcasts. This was a pretty big change for me: I listened to several hours of podcasts a week, often while commuting or at home, and especially while running. My subscriptions were mostly lighter stuff like comedy podcasts (What a Time to be Alive, Blocked Party, and UYD are my three best recommendations) and generally stayed away from more loaded stuff like politics, sports, or anything work-related. It's not like I was flooding my brain with radical messages or important information on a regular basis.

But ever since I've stopped, I've noticed something change about my own writing and thinking. My brain is quieter. It's clearer, and easier to navigate; like the gain on an amplifier had been cranked up for a long time, and we forgot about it, and only when you turn it down do you realize, "Hey that was a lot". I think my weekly writing has gotten better in that period of time, and I've had a few readers email me and ask if I've been doing anything differently. That might be it.

I miss them, but I don't think I'm going to go back. Only after taking them off do you realize that headphones aren't all that good for you. There's a lot of discussion about screen time and how scrolling feeds and glowing screens are having bad effects on us, but a lot less discussion about constant audio stimulation, aside from hearing loss. Maybe this isn't true for everybody; I know a lot of people swear they're more focused and productive with music, for instance, and that's fine. Maybe a hot, hi-fi sensory state is good for some kinds of productivity. Maybe it's just what people like; no more complicated than that.

But I suggest you give it a try. Go for a week with no headphones and no car radio. See how it feels. It's an easier experiment to try than going real cold turkey and putting your phone away for an entire day. See what happens when you turn down the gain on your brain's high-discrimination, hot processing mode. You may not notice anything. But I did.

Permalink to this post is here:

[The Audio Revolution | alexdanco.com](#)

Only a couple links this week since I was busy (sorry) but a few of these are related to one core theme: the disappearing file, and the rise of new kinds of logical structures meant for a streaming internet era. In some ways it's fantastic, in other ways, not so much. I'm going to write about this as its own newsletter topic next week I think, but here are the links now if you want to read em.

[Computer files are going extinct | Simon Pitt](#)

[How Figma's multiplayer technology works | Evan Wallace](#)

[Yahoo Groups is winding down and all content will be permanently removed | Jordan Pearson, Motherboard](#)

And a few others, for good measure:

[Disney, IP, and "Returns to marginal affinity" | Matthew Ball](#)

[New stable form of plutonium discovered | European Synchrotron Radiation Facility](#)

[How two developers made a living with awful games](#) | Ben Reeves

Also this is going to be absolutely awesome, I've been waiting for this for a while:

[Twitch "Watch Parties" let streamers watch Prime Video with viewers](#) | Mariella Moon, Engadget

Have a great week,

Alex