

Voice Cloning Script

Below are four designed-for-voice-capture scripts, each engineered to stress a different part of the speech system. It is meant to be read straight, without acting too hard, but with natural variation.

Designed for phonetic diversity, prosodic variation, emotional contour, pacing shifts, and articulation under different cognitive loads. The goal here is not “interesting content” so much as “linguistic cross-training.”

Recording Notes:

- Record each script as a separate take
- Leave natural pauses
- Do not rush consonants. Do not flatten questions. Pause when you need to breathe, not when punctuation tells you to.
- Read naturally. Do not act. Let the emotional coloration arise from comprehension, not performance
- Drink water between scripts
- Don't re-record unless you make a genuine mistake – slight variability improves generalization.
- Use the same microphone and settings for each recording.
- Keep the microphone in the same relative position to your voice throughout all recordings, don't hold the mic or move yourself around while recording.
- Use a condenser mic and pop filter if you are able.
- Test mic and settings before recording to verify volume, natural reverb, and other conditions match your desired final sound quality.
- Ensure a quiet recording environment. Retake and edit out sections with background noise anomalies.

One

Voice Calibration Script

(General, Use natural speaking or reading voice)

I am going to speak continuously for an extended period of time.

This recording is intentional.

It is not a performance, and it is not a rehearsal.

It is simply a voice, moving through language at a human pace.

At various moments, the rhythm will change.

Sometimes I will speak quickly, stacking clauses and subordinate thoughts.

Other times I will slow down and articulate carefully, allowing consonants to land and vowels to stretch.

This is normal speech.

I am not angry.

I am not excited.

I am not bored.

I am present.

Now let's begin somewhere ordinary.

The room I am in has a temperature.

It is neither cold nor warm enough to demand attention, which is exactly why it works as a starting point.

There is air moving, quietly, probably from a vent that was installed by someone who never expected anyone to notice it.

That person was correct.

I am breathing without thinking about it, which is fortunate, because deliberate breathing is distracting and tends to interfere with speech.

Nonetheless, I can choose to inhale through my nose, exhale through my mouth, pause briefly, and continue.

That pause matters.

Here is a sentence with crisp plosive sounds:

Peter picked a practical pattern of percussive phonemes.

Here is another, slower this time:

Brittle branches broke beneath the boots of bored botanists.

Now softer fricatives:

Sibilant whispers slide smoothly through silent spaces.

And now something uneven:

The value fluctuates irregularly, unexpectedly, inconveniently.

I am not trying to be clever.

I am trying to be complete.

Let's talk about motion.

Walking uphill is different from walking downhill, even when the distance is the same.

Your calves know this.

Your balance system knows this.

Your voice knows this too, in subtle ways, because posture affects airflow, and airflow affects tone.

Imagine describing something while walking.

Now imagine describing the same thing while seated.

The words may be identical, but the voice is not.

I am seated now.

Numbers are useful.

Zero.

One.

Two.

Three.

Four.

Five.

Six.

Seven.

Eight.

Nine.

Ten.

Eleven.

Twelve.

Thirteen.

Fourteen.

Fifteen.

Sixteen.

Seventeen.

Eighteen.

Nineteen.

Twenty.

Thirty.

Forty.

Fifty.

Sixty.

Seventy.

Eighty.

Ninety.

One hundred.

One thousand.

Ten thousand.

One hundred thousand.

One million.

Decimals matter too.

Point one.

Point zero five.

Point nine nine nine.

Fractions sound different.

One half.

Two thirds.

Three quarters.

Percentages shift stress.

Approximately forty-seven percent.

Roughly ninety-two point six percent.

Now a question.

What happens when a question does not sound like a question?

You can hear it, even if the punctuation is missing.

The pitch rises slightly.

The rhythm bends.

And now a statement that pretends to be a question but isn't.

You know exactly what I mean.

Let's pivot.

Imagine explaining something mildly technical to someone intelligent but unfamiliar with the topic.

You slow down.

You define terms.

You avoid jargon until it becomes unavoidable.

For example:

A system is not just a collection of parts.

It is a pattern of relationships.

Remove one component, and the system may persist.

Disrupt the relationships, and it collapses.

That sentence contains confidence, not arrogance.

Those are different sounds.

Now imagine mild disagreement.

I see your point.

I understand the logic.

I still think there's a flaw in the premise.

Notice how that lands without hostility.

Notice the controlled emphasis.

Now, stronger disagreement.

No, that doesn't follow.

The conclusion doesn't match the evidence, and pretending otherwise won't fix it.

Still calm.

Still measured.

Emotion without melodrama.

There is a quiet kind of relief that arrives when a long-running concern resolves itself.

It does not announce its presence.

It simply allows the body to unclench.

There is also quiet frustration.

The kind that doesn't raise its voice.

It just repeats itself internally, over and over, waiting for acknowledgment.

Let's introduce contrast.

Short sentence.

Longer sentence with a dependent clause that wanders slightly before returning to its point.

Very short sentence.

Now something playful.

A purple platypus politely purchased pineapple pizza on principle.

Now something formal.

Pursuant to the prior discussion, the aforementioned considerations remain applicable.

Now something conversational.

Yeah, that's probably not how it's going to go.

Time for lists.

Red, blue, green, yellow, orange, violet.

Circle, square, triangle, rectangle, hexagon.

Wood, metal, plastic, glass, stone.

Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday.

January, February, March, April, May, June, July, August, September, October, November, December.

Proper nouns shift articulation.

Oregon.

Massachusetts.

Albuquerque.

Reykjavík.

Kyoto.

Names matter too.

Alexander.

Beatrice.

Christopher.

Eleanor.

Sebastian.

Say them neutrally.

Say them with warmth.

Say them with mild irritation.

Now a longer, meandering passage.

At some point, every explanation becomes a story, whether it intends to or not.

The moment you choose where to begin, you have already decided what matters.

The moment you decide what to omit, you reveal your priorities.

This is not a flaw in communication.

It is its defining feature.

Language is a tool, but it is also a constraint.

You can feel it when a word almost fits and doesn't.

That hesitation – that micro-pause – is valuable data.

Let it happen.

Let's slow down.

I am speaking more deliberately now.

Each word is allowed to complete itself.

The ends of sentences are not swallowed.

The beginnings are not rushed.

Breath in.

Breath out.

Continue.

Now speed up slightly.

Sometimes speech accelerates because thought accelerates, and sometimes thought accelerates because speech does.

You can feel the feedback loop if you pay attention, especially when explaining something you know well but haven't articulated recently.

That's fine.

End with steadiness.

This recording does not need a conclusion.

It simply needs continuity.

The voice remains consistent even as topics drift.

That consistency is the point.

I am still here.

I am still speaking.

This is still my voice.

Two

Emotional Without Theatrics

(Controlled affect, internal states, tonal contrast)

This script is about emotional shading, not melodrama. Imagine explaining your internal state to someone observant and intelligent who does not need it exaggerated.

I am calm right now.

Not performatively calm, not artificially neutral, just steady.

My breathing is regular.

My jaw is relaxed.

Nothing in particular is demanding urgency.

This is a useful baseline.

Now, mild irritation.

It's the kind that doesn't announce itself.

The kind that appears when a small inefficiency repeats itself unnecessarily.

Nothing is "wrong," exactly – but something is slightly off, and it keeps happening.

That repetition matters.

Now relief.

The moment when a thing you were bracing for simply... doesn't arrive.

Your shoulders drop before you notice them doing it.

Your thoughts widen.

You stop rehearsing explanations you no longer need.

That silence is a sound too.

Now quiet delight.

Not joy with fireworks.

Joy like finding something exactly where you hoped it would be.

Joy that doesn't need witnesses.

Now disappointment.

The understated version.

The one that registers as recalibration rather than collapse.

You adjust expectations.

You move on, but you remember.

Now frustration, but controlled.

I've explained this already.

I know I've explained this already.

The words were clear.

The issue isn't comprehension – it's attention.

Still calm.

Still contained.

Now reassurance.

It's okay.

We have time.

Nothing irreversible has happened yet.

Those sentences land differently when meant sincerely.

Now skepticism.

I hear what you're saying.

I understand the structure of the argument.

I don't think it holds.

Notice the absence of hostility.

Now confidence.

This is familiar territory.

I know how this works.

I've seen the failure modes, and I know which ones matter.

Now fatigue.

Not exhaustion.

Just the awareness that rest would be useful.

That continuing is possible, but optional.

And finally, neutrality again.

I am here.

I am speaking.

Nothing more is required.

Three

Articulation Stress-Test

(syllabic diversity, mouth gymnastics)

This script is intentionally strange. Meaning is secondary. Precision is everything.
Take it slowly. Over-articulate without caricature.

The quick brown fox jumps over the lazy dog.
Pack my box with five dozen liquor jugs.
Sphinx of black quartz, judge my vow.

Now variations.

A brisk, bronze fox vaulted gracefully over lethargic hounds.
Six slick swans swiftly swam southward.
Jaded wizards hex quick nymphs before lunch.

Now consonant clusters.

Strange strength stretched through the stratosphere.
Crisp crusts crackled catastrophically.
Blunt instruments bruised brittle branches.

Plosives.

Battered backpacks bumped bus benches.
Proper paper packages popped open prematurely.

Fricatives.

Soft silver surfaces shimmer subtly.
Feral shadows slid silently sideways.

Liquids and glides.

Rural jurors rarely roar redundantly.
Luminous lilies linger lazily.

Vowels, stretched.

Meeet meee at theeeee oooooold stooooone doooooor.
Ah. Eh. Ee. Ih. Oh. Ooh.

Diphthongs.

Now. Here. Voice. Choice. Layer. Fire.

Nonsense, but pronounceable.

Flindercrash wobulated the praxenfold mechanism.
Quindlex murmured obliquely beneath the varnished trellis.
Zamperish plothkins jittered uneasily.

Now faster.

Plinths, trusses, glyphs, myths, breaths.
Axle, pixel, kernel, vowel, symbol.

Now slow again.

Antidisestablishmentarianism.
Electroencephalographically.
Incomprehensibilities.

Let them land.

End this section calmly.

The mouth rests.
The tongue settles.
The voice continues.

Four

Dialogue-Driven Cadence & Response Inflection

Read both sides aloud. Do not exaggerate the contrast or do “voices” – just let the intent shift.

A: Are you ready?

B: Yes. I think so.

A: That didn't sound convincing.

B: It was honest. Not everything ready feels confident.

A: Fair enough. Do you want to start, or should I?

B: You should. You frame things more clearly.

A: That's generous. Or strategic.

B: Possibly both.

A: Here's the issue. We've been circling the same point for a while.

B: I know. I was hoping repetition would clarify it.

A: Has it?

B: For me? Yes. For you? Apparently not.

A: That's not quite fair.

B: It wasn't meant to be.

A: Then say what you mean.

B: I mean we're using different definitions.

A: Of the same word?

B: Of the same outcome.

A: That explains the tension.

B: It does. It doesn't solve it.

A: No. But it narrows it.

B: Progress, then.

A: Small progress.

B: Still counts.

A: Are you comfortable moving forward like this?

B: Comfortable enough. Comfort isn't the goal.

A: Clarity is?

B: Clarity and momentum. In that order.

A: Then let's continue.

B: Agreed.

At this point, you're past "sufficient."

You've given the model not just your voice, but your behavior in language – which is the difference between a clone and a ventriloquist dummy. This is how you give a machine a real voice instead of a polite impersonation.

Nice work!

For the best voice-modeling, record a total of two hours, including this script, of speaking. Use the voice or tone you'd like your model to best portray (eg. voice-over, audiobook, conversational, monologue). Read from books or other media, or speak clearly and candidly, what you say doesn't matter so much as how you say it. Send me a link to your voice model so I can check it out, thanks!