

The AI Podcast Ad Workflow

You have a script the client already signed off on. This is how we turn it into a finished AI podcast ad with two tools, Claude and Google Omni, and the exact prompts and settings at each step. No camera, no studio, no shoot. About five dollars a video.

SCRIPT TO FINISHED AD, STEP BY STEP

This picks up where the writing ends. You have a script the client is happy with, and now you need it as a finished AI podcast ad: a talking-head video that looks like a clip from a podcast, made without a camera, an actor or a shoot. The whole pipeline runs on two tools. **Claude** preps the script, and **Google Omni** turns a single photo into a host that speaks it. The words do not change. We are only turning them into video.

1. Break the script into clips

Omni's lip-sync holds for about eight seconds, then it drifts. So the first move is to have Claude cut the approved script into single lines, each one a clip of eight seconds or less, without touching the wording. A rough gauge for length is word count divided by three, so around 22 words is your ceiling per clip.

Split this approved script into individual clips for AI video, without changing a single word.

Rules:

- Each clip is one continuous line of speech, 8 seconds max (roughly 22 words)
- Break on natural pauses, never mid-thought
- Keep the wording exactly as approved
- Number each clip and show its word count and rough length (words / 3)

Script:

[PASTE APPROVED SCRIPT]

2. Prep the pronunciation

Omni will mangle brand names, product names and acronyms unless you spell them the way they sound. Have Claude build a respelling list before you render anything. These phonetic spellings go only into the line Omni speaks, never onto the screen.

List every word in this script Omni might mispronounce: brand names, product names, acronyms and unusual spellings.

For each, give a phonetic respelling that reads correctly out loud (for example "Gelpro" becomes "Jelpro"; letter-by-letter acronyms like "NDIS" stay as-is).

These respellings go only in the line Omni speaks, never in the on-screen text.

Script:

[PASTE APPROVED SCRIPT]

3. Lock your host

Pick one clear reference photo. A real selfie, a team member, or a generated character, it does not matter, as long as it is the same face on every clip. That single image sets the entire look of

the ad: the person, the framing, the room, the lighting. Choose it as carefully as you would cast a shoot, because in this workflow it is the casting.

4. Render each clip in Omni

This is where it becomes a video. Feed Omni the one reference photo and a single clip line, and it renders a talking head that speaks the line in a real voice, no separate voice-over step. Do one clip at a time, vertical 9:16, and feed the same photo every time so the host never changes. The gotchas below are the difference between a clean host and a warped mess.

OMNI RENDER BRIEF · ONE PER CLIP

Reference: one clear photo of your host (the SAME photo on every clip)
Line: "[EXACT CLIP LINE, with phonetic respellings]"
Shot: static camera, medium close-up, vertical 9:16
Voice: warm, [accent], high energy, not flat
Duration: match the line, keep the clip under 8 seconds

Gotchas that ruin Omni renders:

- No "podcast desk", "ring light" or "microphone" (renders as warped props)
- No emotion adjectives like happy or excited (distorts the face)
- No motion described before the line (creates a dead, frozen second)
- One line per clip

5. Stitch the clips in order

Lay every rendered clip end to end in script order and join them into the full ad. Keep the frame rate the same across all clips or the join will stutter, and hold the last frame an extra beat so the final line finishes cleanly instead of getting clipped. What you have now is a rough cut of the whole ad, in one file.

6. Verify every frame

This is the step that separates a shippable ad from one that makes you look like you cut a corner. Watch the rough cut back twice, once at normal speed and once frame by frame. Then re-roll

only the clips that fail, never the whole ad.

QC CHECK · BEFORE YOU CAPTION

Watch the full ad back twice: once at normal speed, once frame by frame. Re-roll the clip in Omni if you see:

- A hallucinated logo, label or text anywhere in frame
- A background that warps, slides or morphs
- A dead second before the person starts talking
- Lip-sync drifting off the voice (usually clips over 8s)
- The host's face shifting between clips

Re-roll only the broken clips, then drop the fixes back into the stitch.

7. Caption and ship

Most people watch on mute, so captions are not optional. Burn them across the whole ad, put the opening line on screen as bold text for the first three seconds, export at 9:16, and it is ready to post. From approved script to finished ad, no camera in sight.

The stack, and what it actually costs

Two tools do the work, and the script is already done and paid for. Everything downstream of it is a few dollars and an afternoon.

Stage	What we use	Rough cost
Script	Already approved, your starting point	Done
Clip split and pronunciation	Claude	Cents
Voice and talking-head render	Google Omni	About \$5 per finished ad
Stitch, captions, finish	CapCut or your editor of choice	Free

The short version

Clips in Claude, render in Omni, stitch, verify, caption, post. One approved script becomes a finished podcast ad in an afternoon, and the same script becomes a second, a third and a tenth just by swapping the host photo or the opening clip. That is how one signed-off idea turns into a month of creative.

If turning your approved scripts into fifty finished ads a month sounds better than making one, that is the job we do for brands. [Book a free creative audit](#) and we will show you what it looks like run off your real spend data.

Frequently asked questions

I already have a script. What does this workflow actually do?

It turns that approved script into a finished talking-head video ad using two tools. Claude preps the script, and Google Omni renders a talking head from a single reference photo. Omni generates the voice and the video together, clip by clip. You stitch the clips, check every frame, caption it, and post. The words never change. This is the production half only, not the writing.

Do I need to be on camera or hire an actor?

No. One reference photo plus the approved script becomes the host. That is the whole point: a signed-off script becomes a finished ad without booking a shoot, which is what makes it possible to produce fifty a month instead of one.

How long does it take and what does it cost?

Once the script is signed off, an afternoon. Omni renders the voice and the video together for around five dollars in credits for a 40-second ad. Claude does the text prep for cents. The time goes into the verify pass, not the making.
