

FUTUR BLOOF

SAY HER NAME

PROMPTING A MASTERPIECE — process notes



[Watch it · YouTube](#)

WHAT IT IS

A 17-second drop. Neon alley, heavy rain, a ronin holding a blade at a kneeling man's throat, and a question he does not want to answer. Three shots, MiniMax H3, no compositing. The interesting part is that the first version of this film did not work at all, and fixing it produced a rule about film grammar that no amount of prompt engineering would have found.

THE PROBLEM: YOU CANNOT CONTINUE RAIN

The original plan was one continuous scene split across two generations, with the second continuing the first. Standard technique: bind the tail of shot one as a video reference so the second picks up where it left off.

It failed, and it failed in a way worth understanding. The characters held. The poses held. The lighting held. But the RAIN restarted — a completely different pattern of falling water — and a holographic sign in the background jumped position. A still or a short clip used as a reference carries POSE, not MOTION FIELDS. In any scene with weather, smoke, crowds or particles, the field re-rolls and no amount of post can hide it.

- A 0.5-second tail as a video reference is below the node's 2-to-15-second contract. It starves silently and the render falls back to prose only — no error, no warning.
- Binding the tail's full audio made the new take mimic the previous DELIVERY almost exactly — the actor stopped acting and started impersonating. Strip speech from the reference bed and pass identity through a separate short voice reference.
- With the bed stripped the delivery came back, but the room tone then stepped at the cut. Each fix broke what the other repaired. That was the signal that the approach itself was wrong.

The cure was not a better reference. It was coverage — cut away instead of continuing.

THE FIX: CUT TO COVERAGE

Instead of continuing the wide, cut to a close-up of the ronin, then to a reverse of the kneeling man. Every new framing re-establishes its own rain, and the eye forgives texture across a shot change. Cinema has been hiding continuity breaks behind reverses for a century; generative video did not repeat that.

Keep eyelines consistent across the cuts — the wide has the ronin facing screen right and the kneeling man screen left, so the close-up and reverse inherit those directions and the three shots read as one conversation.



The close-up. Identity cast from a frame of the wide, so it is the same man.

THE LINE THAT FIXED THE PERFORMANCE

The close-up's dialogue was tagged as a shout. It rendered quieter than its own rain — the line peaked below the ambient bed. Performance tags direct; they do not guarantee.

The fix was typography, not tagging. Writing the pacing INTO the words:

```
before: <d>[English] <pant> <shouts> SAY HER NAME. </d>  
after: <d>[English] <pant> <shouts> SAY... HER... NAME. </d>
```

Same seed, same refs, only the ellipses changed: the line went from 22 dB BELOW its rain bed to 27 dB above it.

The dialogue head performs what is WRITTEN. Beats, pauses and force live in the text's own rhythm, not in the tag. Write the delivery you want to hear.

TRANSCRIBE EVERY TAKE

A performance tag placed after the final word was SPOKEN ALOUD on one roll — the actor said the stage direction. It transcribed as 'Mrs. Whiskers, catch breath'. The same tag performed silently in a different take of the same scene, so this is per-roll, not a rule.

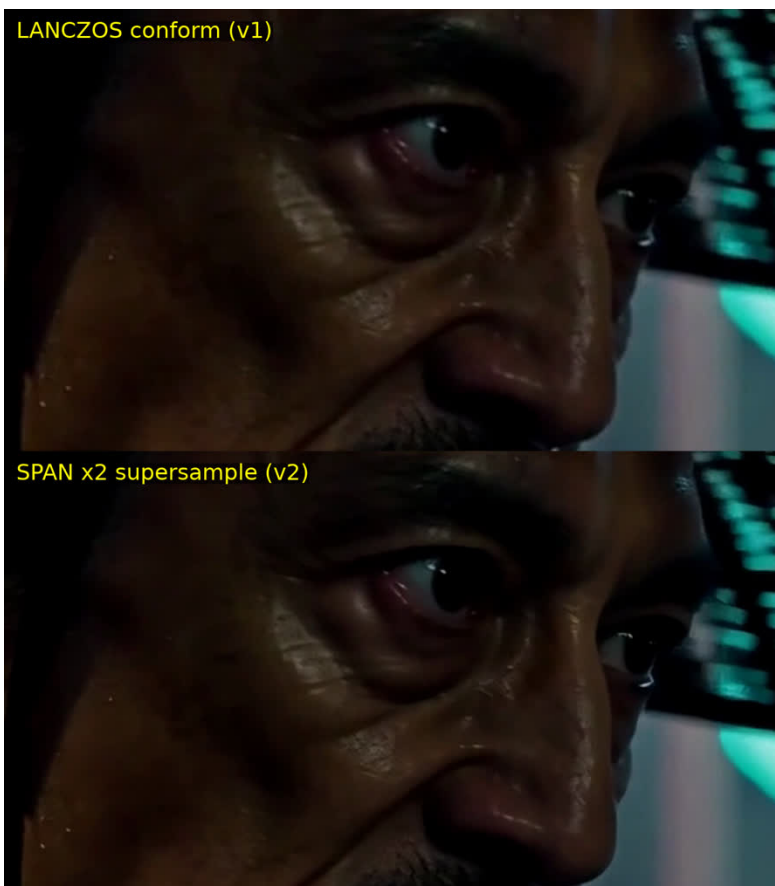
— Run speech-to-text on every take that has dialogue. It takes seconds and it caught this before playback did.

- Put breaths and reactions that follow the last spoken word into the BEAT PROSE, not into the dialogue block.

FINISHING: UPSCALE, DO NOT STRETCH

A conform pass that scales to 1080p is a floor, not an upscale — it satisfies a resolution gate by stretching pixels. Real resolution work is a model upscale past the target followed by a supersample back down.

```
# SPAN 2x through the render queue, then supersample down to the ship target
ffmpeg -i span2x.mp4 -vf "scale=1920:1072:flags=lanczos,pad=1920:1080:0:4:black" \
-c:v libx264 -crf 14 -preset slow -pix_fmt yuv420p -c:a copy master_1080.mp4
```



Same frame, 1:1 crop. Top: lanczos conform. Bottom: SPAN 2x supersampled down. Eyelashes become individual hairs; skin gets pore texture instead of gradient.

One rule that costs nothing: never run a model upscaler over a found object. The end card in this drop is a crop of a real artefact, and it gets plain lanczos to match the container. Enhancing it would be restyling something whose whole value is that it was found.