



An Editor's Guide to Pacing

The *Cut*

Where to cut, how to pace it, and how to know when it is finished. What to do with the footage once you have it.

Written by Ace
Filmmaker and Creative Director

START HERE

Editing is removal.

Almost every piece of advice about editing treats it as a place to add things. It is not. It is the place where you take things away until only the part worth watching is left.

A talking head is the hardest thing to edit well, because there is nothing to hide behind. No cutaways, no B roll, no location. Just a person and whatever rhythm you give them. That rhythm is the whole job, and it is made almost entirely of decisions about what to delete.

WHAT THIS GUIDE WILL NOT DO

It will not tell you what the piece should be about, or whether it was worth making. Editing cannot rescue a video that had nothing to say, and no amount of pacing will fix it. That is a different problem with its own system, and it is not in a craft guide.

The order to work in

Do these in sequence. Doing them out of order is the single biggest waste of time in editing, because you end up polishing material you are about to delete.

PASS	WHAT YOU ARE DOING	DO NOT YET
1. Select	Pick the take. Watch them muted, choose one.	Cut anything
2. Strip	Remove the top, the tail and the dead air.	Think about rhythm
3. Tighten	Cut inside sentences. Find the pace.	Add music or captions
4. Sound	Levels, then music if it earns its place.	Add captions
5. Captions	Burn them in, check the safe area.	Re-cut anything
6. Watch	On a phone, with sound off, then on.	Post it the same hour

ONE OF FIVE · THE STRIP

Cut the first three seconds. Always.

There is a version of your video hiding inside the file, and it starts later than the one you recorded.

Everybody records a runway: settling, a breath, a small adjustment, then the actual first word. All of it has to go. Find the first syllable that carries meaning and cut immediately before it, not a beat before it. A beat before it is still runway.

What to remove, in order

REMOVE	WHY	TYPICAL SAVING
The runway at the top	It is you arriving, not you talking	2 to 5 seconds
The tail after the last word	You reaching for the phone	1 to 3 seconds
Gaps between sentences	Felt right while speaking, reads as slow	Several seconds
Every 'so', 'basically', 'you know'	They carry nothing and they stack up	Often 10 per cent
Restarts and false starts	You already said it better later	Varies
The sentence explaining the next sentence	Almost always unnecessary	Varies

A ninety second recording usually contains sixty seconds of video. That is not a sign you spoke badly. It is what the medium costs, and every practised person you have watched has paid it.

TWO OF FIVE · THE JOIN

Cut on the movement.

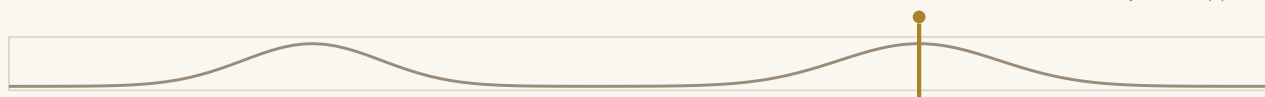
Two cuts can remove exactly the same words and one of them will be invisible while the other makes the viewer flinch.

CUT IN THE STILL PATCH

The join is visible

**CUT ON THE MOVEMENT**

The join disappears



The line is how much you are moving. Peaks are gestures, flats are stillness.

The difference is what your body was doing at the moment of the join. Cut while you are still and the jump in position is the most interesting thing on screen. Cut while your hand is mid gesture and the eye is already following motion, so it absorbs the change without registering it. This is the oldest trick in editing and it costs nothing to use.

What a bad cut actually looks like

The viewer does not see a cut. They see you teleport. Your head was here and now it is two inches left, your shoulders dropped, and the light on your face changed slightly. None of that registers as an edit. It registers as something being wrong, and the response is to stop watching without ever knowing why.

This is why the fix is placement rather than decoration. You are not trying to make the cut prettier. You are trying to put it somewhere the eye is too busy to notice it.

JUMP CUTS ARE NOT THE ENEMY

A visible jump cut is an accepted convention now, and a video full of them does not look broken to anyone under about forty. The problem is never that a cut is visible. It is that a cut is jarring, which is a different thing. Placed on motion, you can make fifty of them and the piece still flows.

THE JOIN

Three ways to hide a cut

METHOD	HOW	WHEN TO USE IT
Cut on motion	Join while a hand or the head is moving	The default. Works most of the time.
Cut on the breath	Join on the inhale before a new sentence	When nothing is moving
Punch in	Scale the second clip up 10 to 15 per cent	When the join refuses to hide

On the punch in

Scaling up slightly at a cut turns a jump into an apparent change of shot, and it reads as deliberate. Two warnings. Keep it under about fifteen per cent or it becomes obvious, and do not do it at every cut, because a video that creeps closer all the way through ends up somewhere absurd. Alternate in and out, or use it only where you need it.

THE ONE THING NOT TO DO

Do not reach for a transition. A cross dissolve, a whip, a zoom blur or anything that came free with the app announces that a cut happened, which is the exact opposite of what you want. A hard cut placed well is invisible. A transition placed well is still a transition.

THREE OF FIVE · RHYTHM

Vary the lengths.

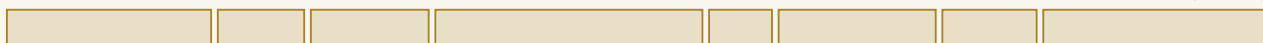
A cut every two seconds is not pacing. It is a metronome, and the brain tunes it out in about fifteen seconds.

EVERY CLIP THE SAME LENGTH

Reads as a metronome

**CLIPS OF DIFFERENT LENGTHS**

Reads as somebody talking



Speech is naturally uneven. Some thoughts take four seconds and some take one, and an edit that respects that sounds like a person while an edit that imposes a fixed interval sounds like a machine. If your timeline looks like a row of identical bricks, you have been cutting to a rhythm instead of to the meaning.

Where to put the long one

Give the most important sentence in the piece the longest uninterrupted run. Cutting into it chops it up and signals that you were bored of your own point. Let it breathe, then cut tightly either side of it so the contrast does the work.

How long is too long

There is no correct number, but there is a useful signal: if you find yourself watching your own hands rather than listening, the clip has run past its usefulness. That moment tends to arrive somewhere between six and ten seconds on a static talking head, and later if you are moving or if the point genuinely needs the room.

IF THE PIECE IS	TYPICAL SHAPE	WATCH OUT FOR
Under 30 seconds	Few cuts. Often one or two.	Cutting out of nervousness rather than need
30 to 90 seconds	Varied, with one long run in the middle	The metronome. This is where it appears.
Over 90 seconds	Needs a change of shot or setting	One unbroken angle for two minutes

The pause you should keep

One. Somewhere in the piece, leave a real silence after a sentence that deserves one. It is the clearest signal available that you meant what you just said, and it is the first thing an over enthusiastic edit removes.

FOUR OF FIVE · SOUND

Sound is most of it.

People will forgive footage that looks average. They will not forgive audio that is hard to listen to, and they leave without being able to say why.

The three that matter

PROBLEM	WHAT IT SOUNDS LIKE	THE FIX
Room echo	Hollow, like a bathroom	Soft things in the room. A bed, curtains, a rug.
Uneven level	Loud, then suddenly quiet	Compression, then one master level pass
Background hum	A constant layer under the voice	Noise reduction, gently. Too much sounds underwater.

Cut the sound and the picture separately

This is the single technique that separates edits that feel professional from edits that do not. Let the audio from the next clip start a fraction before the picture changes, or let the previous audio run a moment past the visual cut. The join stops landing in one place and the whole thing smooths out. Most editing apps let you drag the audio and video edges independently once you unlink them.

ON MUSIC

Music underneath a talking head is optional and usually makes things worse. If you use it, put it far enough down that you would not notice it was there until it stopped, and choose something without a strong beat, because a beat fights your speech rhythm and your speech loses. If the video only works with music, the problem is the video.

FIVE OF FIVE · CAPTIONS

Captions, and where they can live.

Most people will meet your video on mute. Captions are not an accessibility extra, they are the primary text of the piece.



- **MIDDLE BAND ONLY**
Everything outside it gets covered by the app.
- **LOW, NOT CENTRED**
Centre sits on your face. Lower third is clear.
- **TWO LINES MAXIMUM**
Three lines is a paragraph and nobody reads it.
- **BURN THEM IN**
Auto captions are not guaranteed to be on.
- **CHECK ON A PHONE**
Not on the laptop you edited it on.

The frame you export is not the frame anyone sees. Every platform lays its own interface over the top: a caption, your handle, a row of buttons down one side, and a progress bar along the bottom. Anything you place under that furniture is gone, and you will not discover it in your editing app because your editing app shows none of it.

ROUGHLY WHERE	WHAT SITS THERE	SO
Top eighth	Status bar, and sometimes a header	Keep it empty
Bottom fifth	Caption text, handle, progress bar	Keep it empty
Right edge	The column of buttons	Nothing important there
The middle	Yours	Captions low inside it

Approximate on purpose. The exact zones move whenever an app is updated, so build in margin rather than measuring to the pixel.

CAPTIONS

Making them readable

- Two lines, never three. Three lines is a paragraph, and a paragraph on a moving image does not get read.
- Match the spoken word. Captions that paraphrase are distracting, because the viewer hears one thing and reads another.
- Plain and heavy. A bold sans face with a solid backing or a hard outline. Thin type over a moving image is unreadable.
- No animation per word. Words popping in one at a time force the eye to chase them, which is fine for a novelty and exhausting for anything longer than a few seconds.
- Correct the auto captions. They will misspell names and anything technical, and those are exactly the words that matter.

How long each card should stay up

Long enough to read twice at a glance. In practice that means roughly a second and a half for a short line and around three for a full two line card. Captions that change on every few words are the most common fault, and they are exhausting in a way viewers experience as the video being tiring rather than the captions being wrong.

CARD	HOLD IT FOR	THE FAULT
Three or four words	About 1.5 seconds	Flicking past before it is read
One full line	About 2 seconds	Changing mid sentence for no reason
Two full lines	About 3 seconds	Holding it after the speech has moved on

The mute test

Watch the finished piece on a phone with the sound off and read only the captions. If you cannot follow it, neither can the large share of people who will never turn the sound on. This takes ninety seconds and catches more problems than any other single check in this guide.

KNOWING WHEN TO STOP

The signs you have gone too far.

There is a point past which every further cut makes the piece slightly worse, and it arrives earlier than most people think.

Over-editing does not look like a mess. It looks slick and feels airless. Every pause is gone, every breath is gone, and the person on screen sounds like they never stopped to think, which is not a compliment. Watch for these.

- There is no silence anywhere in it.
- You cannot hear a single breath in ninety seconds.
- The cuts have started arriving at a regular interval.
- You have watched it so many times you can no longer tell.
- You are now cutting things because you can, not because they are dead.

THE TEST

Leave it overnight and watch it once in the morning, start to finish, without touching anything. Almost every problem worth fixing announces itself in that single viewing, and almost everything you were agonising over at midnight turns out not to matter.

THE RULE

Finished is when removing one more thing would take something out that the piece needs. Not when it is perfect. Perfect is not available, and chasing it is the main reason people stop posting.

When you are stuck

Being stuck in an edit is almost always one of three things, and none of them are solved by continuing to fiddle with it.

IT FEELS LIKE	IT USUALLY IS	WHAT TO DO
Something is off and I cannot find it	You have watched it too many times	Stop. Come back tomorrow.
It will not tighten any further	It is already finished	Export it and move on
No edit makes this work	The take was not good enough	Reshoot it. It is faster.

The third is the one people resist most and it is the cheapest of the three. Reshooting ninety seconds takes ten minutes. Trying to rescue a flat take can absorb an entire evening and still not work.

REFERENCE

What it looks like, and what caused it.

Symptom on the left. The cause is almost never where you were looking.

WHAT YOU NOTICE	WHAT IS ACTUALLY WRONG	THE FIX
It drags	Gaps between sentences are still there	Strip the dead air first
The cuts are jarring	Joined during stillness	Move the cut onto a gesture
It feels relentless	Every clip is the same length	Vary them. Let one run long.
It starts slowly	The runway is still attached	Cut to the first meaningful syllable
It sounds amateur	Audio, not picture	Level pass, then gentle noise reduction
It ends abruptly	You cut on the last word	Leave a short beat after it
Nobody watches past the start	Usually the first two seconds	Start already talking. See The Take.
It feels airless	Over-edited. Every pause removed.	Put two breaths back in
Captions get cut off	They sit outside the safe band	Move them into the middle third

IN THE APP

The passes, in CapCut.

Everything so far works in any editor. This is the same order of work done in CapCut, the editor I use, on the phone and on a computer.

PASS	ON THE PHONE	ON A COMPUTER
Select	New project, add every take. Watch them with the phone on silent and keep one.	Media, then Import, then drag the take you keep onto the timeline.
Strip and tighten	Move the playhead to the cut, tap the scissors (or long-press, then Split), select the piece you do not want and delete it.	Playhead on the cut, then Ctrl+B (Windows) or Cmd+B (Mac) to split. Delete the piece.
Sound	Tap the clip, swipe the toolbar to Extract audio. The voice becomes its own track. Then Volume for level and Reduce noise for hum.	Select the clip, open the Audio tab on the right, switch on Reduce noise and keep the intensity low.
Captions	Text, then Auto captions, choose the audio source. To fix words, tap a caption, then Edit captions.	Text, then Auto captions, choose the language, then Recognize. Double-click a caption to correct it. Styling is in the Style tab.
Punch in	Select the second clip and spread two fingers on the preview to scale it up 10 to 15 per cent.	Select the clip, open Basic, raise Scale by 10 to 15 per cent.
Export	Tap the resolution label next to Export and choose 1080P, then Export.	Export, top right. 1080p and 30 fps is what CapCut recommends for Reels.

Checked against CapCut's own help pages in September 2026. Button names move between updates, and some features differ between the free and paid versions and between countries. If a button is not where this says, select the clip and look along the toolbar underneath it.

Five habits that save the most time

- Extract the audio before adding B roll. The voice then runs straight through underneath the footage, and you can let the sound lead the picture at a cut, as on the sound page.
- Fix captions as a list. Edit captions shows every line together. Read it once, top to bottom, and correct names and technical words.
- Split a caption that runs long at its natural pause, exactly as you would a clip. It is the quickest way to two lines per card.
- Reduce noise only where you hear hum. On a clean recording it just thins the voice.
- Export at 1080p, not 4K. It is the size CapCut recommends for Reels, and it uploads faster. Then watch it on your phone before posting.

WHAT NOT TO TOUCH

Effects, transitions, templates and animated caption styles. The table above is the whole toolkit. If you are scrolling the effects menu, go back to the rhythm page instead.

PRACTICE

Five drills.

Use footage you already have. None of these are for posting.

01

The strip

DEAD AIR · 10 MINUTES

Take one unedited take and remove only the runway, the tail and the gaps between sentences. Change nothing else. Note the duration before and after.

START WITH One take, straight off the phone

DO Top, tail, and gaps only

DO NOT Cut inside a sentence yet

YOU HAVE IT WHEN It is at least twenty per cent shorter

The mistake everyone makes: Tightening inside sentences on the same pass, which makes it impossible to tell what the stripping alone achieved.

RUN IT IN ONE SITTING

02

Same cut, two places

THE JOIN · 10 MINUTES

Find a cut you have already made. Make a copy and move the join onto the nearest gesture. Watch both. The difference is the whole lesson.

START WITH Any edit with a visible jump

DO Two versions, same words removed

WATCH FOR Which one you notice the cut in

YOU HAVE IT WHEN You can predict it before watching

The mistake everyone makes: Moving the cut so far that the words change, which makes the two versions no longer comparable.

RUN IT IN ONE SITTING

03

Look at the bricks

RHYTHM • 5 MINUTES

Open a finished edit and just look at the timeline. Do not play it. If the clips are all roughly the same width, you cut to a rhythm rather than to the meaning.

START WITH Anything you have already finished

DO Look only at the shape of the timeline

WATCH FOR A row of identical blocks

YOU HAVE IT You can spot it without playing anything

WHEN

The mistake everyone makes: Fixing it immediately. Look at three finished edits first, so you learn to see the pattern rather than one instance of it.

RUN IT IN ONE SITTING

04

Split the audio

SOUND • 10 MINUTES

Take three cuts and pull the audio edge a few frames earlier than the picture edge on each. Watch it back and try to hear what changed.

START WITH A short edit with three or more cuts

DO Unlink audio and video, then offset

WATCH FOR The joins stopping being noticeable

YOU HAVE IT You do it without thinking about it

WHEN

The mistake everyone makes: Offsetting so far that lips and voice separate, which is a different and much worse problem.

RUN IT IN ONE SITTING

05

The mute test

CAPTIONS • 5 MINUTES

Send a finished piece to your own phone. Watch it muted, reading only the captions. Then watch with sound and see what you missed.

START WITH Something you consider finished

DO Watch on a phone, not the machine you edited on

WATCH FOR Captions covered by the app's own interface

YOU HAVE IT Muted and unmuted tell the same story

WHEN

The mistake everyone makes: Testing it in the editing app's preview window, which shows none of the interface that will sit on top of it.

RUN IT IN ONE SITTING



The line this guide will not cross

This is how to assemble it.

Not whether it was worth making.

Everything in this guide is mechanics. Where a cut belongs, how long a clip should run, what to do with sound and captions, and when to stop. All of it can be written down, so all of it has been.

None of it will save a piece that had nothing to say. Editing makes a good idea watchable and makes an empty one shorter, and no amount of pacing changes which of those you are holding. Working out what is worth making, and for whom, needs a system for deciding, and ideally somebody building it with you and being honest about it.

There are two ways to get it. The Feature is the whole system without me: positioning, planning, filming, carousels and an AI setup, in one playbook and a Notion workspace. It opens soon, and because you downloaded this, you will hear first.

The ninety day mentorship is that system built with me, one to one. A filmmaker and creative director in your corner for ninety days: shaping your positioning with you, building your content with you, reviewing your work before it goes out, holding you to the standard and the schedule, answering whatever comes up, and changing course with you when the numbers say so. It is at cinematicauthoritysystem.com. If you want neither, everything above still works.



The Cut

Guide three of three.

The First Frame is the setup. The Take is the performance.

This one is the assembly.

cinematicauthoritysystem.com/guides

[@acethetics.29](#)