

Flujo de trabajo: montaje de sonido directo en Pro Tools.

# Cuando el *timecode* falla: inteligencia artificial y estrategias alternativas para la sincronización en documentales

9

When *Timecode* Fails: Artificial  
Intelligence and Alternative Strategies for  
Synchronization in Documentaries

**Víctor Hugo Herrera Recalde**

Universidad de las Artes (Guayaquil, Ecuador)

[victor.herrera@uartes.edu.ec](mailto:victor.herrera@uartes.edu.ec)

[sound@vicherrera.net](mailto:sound@vicherrera.net)

ORCID: <https://orcid.org/0009-0004-6479-699X>

## RESUMEN

La sincronización entre audio y video es un aspecto crucial en la postproducción de documentales. A diferencia de las producciones guionizadas, los documentales se filman de manera espontánea en múltiples locaciones con diversos equipos, lo que genera importantes desafíos de sincronización. Este artículo explora los principales problemas encontrados al sincronizar audio y video en un documental, comparando los métodos de sincronización basados en forma de onda frente a los basados en código de tiempo. Además, se proponen recomendaciones para establecer un flujo de trabajo eficiente que mitigue estos inconvenientes en futuras producciones.

Palabras clave: sincronización audiovisual, postproducción, documentales, código de tiempo, flujo de trabajo

## ABSTRACT

Audio and video synchronization is a crucial aspect of documentary post-production. Unlike scripted productions, documentaries are often filmed spontaneously in various locations using diverse equipment, which leads to significant synchronization challenges. This article explores the main issues encountered when synchronizing audio and video in a documentary, comparing waveform-based synchronization methods with timecode-based approaches. Furthermore, we propose recommendations to establish an efficient workflow that mitigates these challenges in future productions.

Keywords: audio/video synchronization, post-production, documentary, timecode, workflow

## Introducción

Los documentales se caracterizan por procesos de rodaje abiertos, no repetibles y frecuentemente condicionados por factores externos. Esta naturaleza flexible, aunque esencial para su valor narrativo, introduce desafíos significativos en la etapa de postproducción, particularmente en la sincronización entre audio e imagen. Un desfase perceptible puede afectar la credibilidad del relato audiovisual y la experiencia del espectador.

El presente estudio examina los problemas de sincronización detectados en un documental filmado en cuatro ubicaciones —India, Ecuador, Estados Unidos (Arizona) y Zimbabue— utilizando configuraciones de grabación diversas. La ausencia de un protocolo técnico unificado durante el rodaje generó discrepancias temporales entre las señales de audio y video, evidenciando la necesidad de estrategias robustas de sincronización en flujos documentales contemporáneos.

Para el análisis se emplearon DaVinci Resolve Studio como editor no lineal (NLE) y Pro Tools como estación de trabajo de audio digital (DAW). Aunque estas herramientas permiten gestionar grandes volúmenes de material, la heterogeneidad del flujo de trabajo obligó a implementar soluciones adicionales que se detallan a lo largo del artículo.

11

## Fundamentos técnicos de la sincronización audiovisual

La sincronización audiovisual se fundamenta en la correspondencia temporal entre un evento visual y su manifestación sonora. En flujos de trabajo audiovisuales profesionales, esta alineación es un requisito técnico esencial para garantizar la coherencia narrativa y perceptual del material final. No obstante, la relación entre sincronía técnica y percepción humana no siempre es directa, ya que la experiencia del espectador está mediada por procesos cognitivos y contextuales.

Eg y Behne señalan que «los seres humanos pueden tolerar asincronías audiovisuales que superan los 200 milisegundos»<sup>1</sup>. Esto explica por qué determinados desfases temporales pueden pasar inadvertidos para el espectador aun cuando existe una desalineación técnica objetiva.

En el ámbito cinematográfico y televisivo es habitual el uso de sistemas de doble grabación en los que la imagen se registra en cámara y el audio en dispositivos externos de mayor calidad. Este procedimiento mejora el resultado sonoro, pero exige mecanismos precisos de sincronización en postproducción. Tradicionalmente, la claqueta ha cumplido esta función como referencia audiovisual compartida. Sin embargo, los flujos contemporáneos incorporan métodos automatizados basados en análisis de forma de onda o en el uso de código de tiempo (timecode), conforme a estándares profesionales definidos por la Society of Motion Picture and Television Engineers (SMPTE).

Según la norma de la SMPTE, ST 12-1:2014, «el ‘Time and Control Code’ no es solo un reloj, sino un sistema de etiquetado diseñado primordialmente para identificar cuadros (frames) discretos o pares de cuadros, permitiendo establecer una relación temporal precisa entre diversas fuentes de señal»<sup>2</sup>.

12

## Captura multicanal en documentales

Los documentales suelen filmarse en condiciones impredecibles, con equipos diversos y sin las comodidades de un set controlado. A diferencia de una película de ficción (donde se puede repetir una toma si algo falla), en un documental uno puede estar capturando momentos únicos e irrepetibles en múltiples locaciones. Es común que se utilicen varias cámaras

---

1 Ragnhild Eg y Dawn M. Behne, «Perceived Synchrony for Realistic and Dynamic Audiovisual Events», *Frontiers in Psychology* 6, n.º 736 (2015): 1-12, <https://doi.org/10.3389/fpsyg.2015.00736>.

2 Society of Motion Picture and Television Engineers, «SMPTE ST 12-1:2014-Time and Control Code. White Plains», SMPTE, 2014.

a la vez, drones, grabadoras de sonido portátiles, micrófonos inalámbricos (*lavaliers*), etc., cada uno quizás manejado por una persona diferente, o a veces por un solo operador. Todo esto genera un escenario multicanal complejo: múltiples fuentes de audio y video que luego deben sincronizarse para contar una historia coherente.

En producciones audiovisuales, «el uso de grabadoras de audio independientes del sistema de cámara —conocido como doble sistema— requiere un método preciso para mantener la sincronización entre imagen y sonido durante la postproducción»<sup>3</sup>.

Veamos algunos ejemplos típicos:

- Entrevistas: el caso más frecuente es una entrevista en la que la cámara graba imagen con un micrófono de referencia, pero el audio principal proviene de un grabador externo con micrófonos de mayor calidad (por ejemplo, un *boom* o micrófono de corbata conectado a una grabadora Zoom o Tascam). Aquí tenemos al menos dos pistas de audio por toma (cámara + grabadora) que deben alinearse después. Si se usan dos cámaras para diferentes ángulos, el material crece a cuatro fuentes (dos videos, dos audios).
- En un rodaje estilo reportaje: el camarógrafo puede llevar un micrófono *shotgun* montado en cámara mientras el sonidista captura en paralelo con una grabadora portátil todo lo que ocurre. Puede haber momentos en que la cámara deje de grabar (para cambiar batería, tarjeta, etc.) pero el sonido siga corriendo continuamente para no perder detalle. Esto da lugar a clips de audio continuos muy largos y múltiples clips de video más cortos intercalados.
- Multicámara en eventos: para cubrir un evento real (una boda, un concierto, una manifestación) un documentalista puede desplegar

---

<sup>3</sup> David Miles Huber, Robert E. Runstein y Richard Holmes, *Audio Post Production for Television and Film: An Introduction to Technology and Techniques* (Oxford: Focal Press, 2004).

varias cámaras en distintos puntos y un equipo de sonido central. Es probable que no todas las cámaras tengan capacidad de recibir código de tiempo (por supuesto, muchas podrían ser DSLR o cámaras de bolsillo). Por tanto, algunas grabarán con *timecode* y otras solo con audio guía. En postproducción tocará emparejarlo todo manual o semiautomáticamente.

- Producción «solo-crew»: en producciones documentales independientes es frecuente la figura del realizador que asume simultáneamente los roles de dirección, cámara y registro sonoro. En este tipo de configuraciones, un solo operador puede trabajar con múltiples cámaras —por ejemplo, una fija en trípode y otra en cámara en mano— junto con varios micrófonos inalámbricos conectados a grabadoras portátiles que permanecen registrando audio de forma continua durante largas jornadas.

14

Este esquema operativo dificulta la activación y detención simultánea de todos los dispositivos, por lo que resulta habitual que los micrófonos *lavalier* graben de manera ininterrumpida mientras las cámaras se inician y detienen de forma intermitente según las acciones relevantes del rodaje. Como consecuencia, al sincronizar posteriormente por forma de onda, los fragmentos de audio continuo no siempre presentan una correspondencia directa con los clips de video, ya que el audio puede comenzar antes o finalizar después del registro visual.

Diversos sistemas de sincronización automática tienden a recortar el audio externo para ajustarlo a la duración del clip de video, lo que puede provocar la pérdida de información sonora relevante registrada fuera de cuadro. Esta situación incrementa el tiempo de búsqueda manual en postproducción y evidencia una de las principales limitaciones de la sincronización basada exclusivamente en forma de onda en contextos documentales sin código de tiempo.

Como vemos, la falta de sincronía instantánea en rodaje (ya sea por no usar claqueta, por no disponer de *timecode* en todos los equipos,

o por parar/arrancar dispositivos a destiempo) deriva en bastante trabajo manual posteriormente. Por eso, incluso en contextos documentales, es recomendable introducir la mayor organización posible durante la filmación:

- Si se cuenta con los recursos, utilizar código de tiempo en todos los dispositivos es lo ideal. Herramientas como transmisores y receptores de timecode —por ejemplo, sistemas Tentacle Sync o Deity TC— permiten realizar *jam-sync* de forma inalámbrica entre cámaras y grabadoras, garantizando una referencia temporal común. Esta práctica es ampliamente adoptada en rodajes documentales complejos, ya que proporciona un marco temporal compartido incluso en configuraciones con múltiples cámaras, drones o dispositivos portátiles, lo que facilita la coherencia narrativa y la eficiencia en la postproducción.
- Si no es posible usar *timecode* (por limitaciones de equipo o presupuesto), se recomienda al menos grabar audio guía en cada cámara. Puede ser el micrófono integrado o un receptor inalámbrico que envíe una mezcla del sonidista a la cámara. Así, aunque no tengan relojes sincronizados, todas las grabaciones tendrán una pista de audio en común que servirá para alinearlas por forma de onda. En otras palabras, nunca dejar una cámara rodando totalmente muda, porque después no habrá por dónde «agarrarla» para sincronizar.
- La generación de referencias audibles o visuales durante el rodaje constituye una estrategia complementaria para mantener la coherencia temporal entre las distintas fuentes de registro. Además del uso de la claqueta al inicio de una escena, la producción de sonidos distintivos visibles y audibles en cámara —como palmadas o aplausos— permite crear puntos de referencia distribuidos a lo largo de la jornada de grabación. Estas marcas facilitan la recalibración de la sincronización en postproducción cuando se producen pérdidas temporales de alineación entre audio y video.

- La documentación sistemática durante el rodaje constituye un recurso fundamental para la postproducción audiovisual. Registrar de manera organizada qué cámaras estuvieron asociadas a determinadas fuentes de audio, así como eventuales incidencias técnicas —como pérdidas de sincronía o fallos en el *jam-sync*— permite anticipar problemas y optimizar la toma de decisiones en la etapa de edición. Esta información de campo resulta especialmente valiosa para identificar segmentos que requerirán mayor atención durante los procesos de sincronización y conformado.

En suma, los documentales presentan desafíos particulares de sincronización debido a su naturaleza espontánea y multiequipo. Una buena parte se puede mitigar en la etapa de rodaje con previsión y uso de herramientas (aunque sea una simple claqueta). Aun así, siempre habrá que recurrir a dos métodos principales en la postproducción: la sincronización por forma de onda y la sincronización por código de tiempo, que explicamos a continuación con sus pros y contras.

16

## Desarrollo

### Metodología y contexto de producción

El estudio se basa en un enfoque cualitativo de tipo aplicado, sustentado en el análisis de un caso real de postproducción documental. El proyecto involucró múltiples cámaras, grabadoras portátiles y sistemas de captación sonora operados de manera no simultánea. Esta configuración refleja condiciones habituales en producciones documentales independientes y de presupuesto medio.

El material fue organizado por jornadas de rodaje y procesado inicialmente en DaVinci Resolve, donde se realizaron tareas de ingestión, sincronización preliminar y generación de archivos *proxy*. Posteriormente

te, el montaje fue exportado mediante archivos AAF para su conformado y mezcla en Pro Tools. Las desviaciones detectadas durante este proceso permitieron evaluar la eficacia de distintos métodos de sincronización.

## Métodos de sincronización

Existen dos métodos principales para sincronizar audio y video en la postproducción, los cuales se detallan a continuación:

### 1. Sincronización por forma de onda

La sincronización por forma de onda se basa en la comparación del contenido sonoro registrado por la cámara con el audio capturado por grabadoras externas. Los algoritmos de los NLE identifican patrones comunes —como picos transitorios— para alinear automáticamente las pistas.

Este método resulta accesible y eficaz cuando el audio de referencia es claro y contiene eventos reconocibles. No obstante, su fiabilidad disminuye en entornos ruidosos, con señales degradadas o cuando existen tramos de audio sin correspondencia visual. Además, este sistema no corrige derivas temporales prolongadas, lo que puede requerir ajustes manuales adicionales.

En la práctica, los editores no necesitan ajustar cada toma a mano: la mayoría de los programas de edición no lineal incluyen funciones de *auto-sync* por *waveform*. Estas herramientas analizan las ondas de audio y calculan el desplazamiento correcto. Por ejemplo, en DaVinci Resolve se puede seleccionar un clip de video y su audio externo, y elegir «Auto Align Clips» → «Based on Waveform» para que el *software* los sincronice automáticamente. El algoritmo busca similitudes en las gráficas de audio y ajusta uno respecto al otro.

**Ventajas:** la sincronización por forma de onda es accesible y no requiere ningún equipamiento especial en rodaje; funciona *a posteriori* con cualquier material que tenga audio común. Es eficaz cuando ambas fuentes contienen sonido claro y con poco ruido de fondo, ya que los patrones

de onda coincidentes se detectan con facilidad. Este método ha salvado a muchos proyectos independientes que no contaban con *timecode*.

**Desventajas:** la eficacia cae cuando el audio de cámara es de mala calidad o tiene mucho ruido. En entornos muy ruidosos (una calle concurrida, un concierto) o si el micrófono de referencia captó sonido distorsionado, es posible que el *software* no encuentre coincidencias claras y el editor deba intervenir manualmente. También puede haber problemas si la pista de cámara y la externa no tienen contenido simultáneo (por ejemplo, si una grabadora registró audio en un momento en que la cámara estaba apagada). En esos casos, la forma de onda no ayudará. Además, la sincronización automática por *waveform* suele alinear el comienzo, pero no corrige posibles derivas: si una cámara corrió ligeramente lenta, las ondas se desfazarán progresivamente. Otro factor a considerar es el tiempo de proceso: analizar largos períodos de audio puede ser intensivo, y en proyectos con horas de material el *autosync* puede tardar o incluso fallar. Por eso, a veces es necesario trabajar por partes (día por día, cámara por cámara).

18

**Uso típico:** dado que es un método que ahorra *hardware*, la sincronización por forma de onda es común en producciones de bajo presupuesto, cortometrajes, entrevistas sencillas y documentales pequeños. También es el plan de respaldo cuando el *timecode* falla o está ausente: muchos editores primero intentan sincronizar por *timecode* y, si no funciona, recurren a *waveform*. Hoy en día, con la integración de estas funciones en Premiere, Final Cut, Resolve, etc., herramientas externas como PluralEyes han caído en desuso (de hecho, Red Giant/Maxon discontinuó PluralEyes en 2023, marcando el fin de sus actualizaciones; no obstante, sus capacidades viven en los flujos de trabajo actuales e inspiran nuevas alternativas).

## 2. Sincronización por código de tiempo

El código de tiempo proporciona una referencia temporal común incrustada como metadato en cada archivo audiovisual. Cuando todos los dispositivos comparten el mismo reloj, la alineación en postproducción se realiza de forma automática y precisa a nivel de fotograma.

Este método es el estándar en flujos profesionales debido a su eficiencia y precisión. Sin embargo, su implementación exige planificación, equipamiento compatible y verificación constante durante el rodaje. Fallos en la configuración o pérdida de sincronía pueden invalidar el sistema, obligando a recurrir a métodos alternativos.

**Ventajas:** cuando se configura correctamente en set, el *timecode* garantiza una sincronización muy precisa a nivel de fotograma. En flujos de trabajo profesionales es el método más eficiente porque prácticamente elimina el trabajo manual. Como comenta Henri Rapp (ingeniero de sonido): «El código de tiempo es el método preferido para alinear audio y video en entornos profesionales, ya que requiere mucho menos trabajo que otros métodos». Con *timecode*, uno puede sincronizar automáticamente cientos de horas de material en segundos, algo inviable manualmente. Es especialmente poderoso en documentales con muchas cámaras y largos periodos de grabación, donde otras técnicas se vuelven lentas o poco fiables. Además, el uso de *timecode* permite mantener la referencia temporal en todo el flujo: por ejemplo, un editor puede montar en vídeo y luego entregar una secuencia a un mezclador de audio, y este podrá cargar los archivos originales en Pro Tools Ultimate mediante el Field Recorder Workflow, utilizando el *timecode* para reconectar de forma instantánea y precisa todo el material en su posición correspondiente.

**Desventajas:** la principal desventaja es que exige planificación y equipo en la etapa de producción. Todos los dispositivos deben ser compatibles con código de tiempo. Configurar varios equipos con *timecode* conlleva costo y complejidad: se necesitan generadores/relojes maestros, cables o transmisores, y asegurar baterías y memoria extra para correrlos continuamente. Para un operador en solitario o un equipo pequeño, esto puede ser engorroso. Como vimos, aun con todo el sistema sincronizado, pueden surgir desviaciones mínimas por la deriva de los relojes internos, en especial si no se resincronizan periódicamente durante rodajes prolongados. Al respecto, Huber *et al.* advierten que «los sistemas esclavos de sincronización tienden a presentar pequeños retrasos, que

pueden oscilar desde fracciones de cuadro hasta un cuadro completo, especialmente en configuraciones prolongadas»<sup>4</sup>. Es decir, el *timecode* no es infalible: puede haber discrepancias de un fotograma aquí o allá, o problemas si alguien olvida activarlo en algún aparato. Otro inconveniente es que no todas las cámaras de bajo costo lo soportan; integrar una DSLR o *smartphone* a un rodaje con *timecode* requiere soluciones creativas (como alimentarle una señal LTC por el micrófono). En resumen, el código de tiempo agrega una capa de robustez, pero a cambio de exigir rigurosidad: todos los equipos deben sincronizarse desde el inicio y recalibrarse en jornadas largas; si esto falla, se pierde el beneficio y toca recurrir al oído.

20

En producciones audiovisuales profesionales y de gran escala, la sincronización mediante código de tiempo constituye el estándar operativo. Estos entornos suelen adoptar esquemas de redundancia técnica que combinan múltiples referencias temporales, como el uso de *timecode* y *genlock* en cámaras, *timecode* y *word clock* en grabadoras de audio, así como claquetas electrónicas con visualización de código de tiempo y señal de audio guía enviada a cámara.

En contextos documentales, la disponibilidad de referencias temporales claras resulta determinante para la correcta alineación entre audio e imagen. La ausencia de eventos visuales recurrentes o de marcas audibles consistentes puede dificultar la detección de desincronías, especialmente cuando el registro sonoro es continuo y la imagen se activa de forma intermitente.

En este sentido, Eg y Behne afirman que «las secuencias de mayor duración contienen múltiples eventos correspondientes que ofrecen pistas temporales audiovisuales»<sup>5</sup>, facilitando la percepción de sincronía en contextos complejos.

Desde esta perspectiva, la generación deliberada de referencias audibles o visuales durante el rodaje puede entenderse como una estra-

---

4 Huber, Runstein y Holmes, *Audio Post Production...*

5 Eg y Behne, «Perceived Synchrony...».

tegia para incrementar la cantidad de pistas temporales disponibles en postproducción. Prácticas como el uso de claquetas, sonidos guía o eventos audibles visibles en cámara adquieren así un fundamento perceptual que complementa su utilidad técnica. Para recapitular las diferencias, presentamos una tabla que contrasta ambos métodos:

**Tabla 1.** Comparativa: forma de onda vs. código de tiempo

criterio	Sincronización por forma de onda	Sincronización por código de tiempo
Precisión	Depende de la claridad del audio; suele lograrse precisión suficiente «al oído», aunque a veces requiere microajustes manuales.	Precisa a nivel de cuadro ( <i>frame</i> ) cuando todos los dispositivos comparten el mismo reloj. La alineación es exacta en cada toma (si no hay deriva de relojes).
Requisitos en rodaje	Que las cámaras graben algún audio común con las grabadoras (por ejemplo, sonido ambiente o claqueta). No necesita <i>hardware</i> especial ni configuración previa de relojes.	Dispositivos equipados con entradas/salidas de <i>timecode</i> o receptores LTC. Todos deben haberse <i>jam</i> -sincronizado antes de filmar y preferiblemente mantenerse sincronizados durante el rodaje.
Rapidez en post	Puede requerir alineación manual en cada toma o confiar en algoritmos automáticos. Con gran volumen de clips, la sincronización por <i>waveform</i> consume tiempo de procesamiento y revisión humana.	Alineación prácticamente instantánea: los NLE ordenan los clips por su marca temporal en uno o dos clics. Ideal para proyectos con cientos de horas, donde ahorra muchísimo trabajo.
Equipo necesario	<i>Software</i> de edición con función de <i>autosync</i> por <i>waveform</i> (Premiere, Final Cut, Resolve, etc.). Opcional: <i>plugins</i> especializados (PluralEyes, etc.) en caso de flujos externos.	Reloj maestro de código de tiempo (puede ser una caja generadora o la clock de una cámara/grabadora principal) y cables o sistemas inalámbricos para distribuirlo. Baterías adicionales para mantenerlos.

Fuente: Elaboración propia.

		En algunos casos, <i>clapperboards</i> electrónicos con <i>display</i> de TC.
<b>Confiabilidad</b>	Funciona bien con audio de referencia limpio y eventos sonoros claros. Puede fallar con ruido excesivo, audio desincronizado o silencios prolongados sin referencias. La verificación y ajustes manuales son comunes.	Muy confiable si se implementa correctamente en set: la sincronización en post es prácticamente automática. No obstante, hay que vigilar posibles derivas de <1 cuadro en tomas largas (mitigable con <i>genlock</i> ) y tener un plan B si algún dispositivo pierde la señal de TC.
<b>Ámbito de uso típico</b>	Producciones independientes, rodajes sin presupuesto para TC, casos donde se olvidó o falló el <i>timecode</i> . También como salvavidas en proyectos profesionales cuando algún clip aislado carece de TC..	Producciones profesionales (cine, TV, docuseries de gran presupuesto, realidad), eventos multicámara, proyectos donde el tiempo de edición apremia. Es el método estándar en entornos donde se prioriza eficiencia en post y se cuenta con equipos compatibles.

## Flujo de trabajo en postproducción (paso a paso)

En un escenario ideal, todo el material de rodaje se organizaría en carpetas separadas por jornada, lo que permitiría sincronizar audio y video mediante código de tiempo de forma rápida y eficaz. Este planteamiento no elimina la necesidad de registrar *scratch* audio en cámara: la ley de Murphy en postproducción recuerda que cualquier dispositivo o sistema de sincronía puede fallar y siempre conviene contar con una pista de referencia. En este sentido, Huber, Runstein y Holmes señalan que «rutar el audio del mezclador tanto a la cámara como a una grabadora externa genera una copia de respaldo automática, crucial cuando surgen problemas

con el material sonoro principal»<sup>6</sup>. Lo óptimo, por tanto, es disponer de un código de tiempo común en todos los dispositivos y, además, de audio guía en cada toma, garantizando la redundancia necesaria para simplificar la sincronización.

Una vez ordenado el material, se importa a DaVinci Resolve para el procesamiento inicial. El flujo básico consiste en ingerir y revisar los clips, sincronizar audio y video —preferentemente con *timecode*, o con forma de onda si este falla— y generar archivos *proxy* de baja resolución que alivian la edición. Tras cerrar el montaje, se exporta un AAF desde DaVinci Resolve, lo que permite iniciar la postproducción de audio en Pro Tools sin perder la referencia a los originales.

En la práctica, este flujo sufrió desviaciones. El editor optó por trabajar en Final Cut Pro con los *proxies* creados en Resolve, en lugar de permanecer dentro de Resolve. Ese cambio complicó la conservación de la metadata necesaria para Pro Tools; fue preciso verificar que Final Cut exportara correctamente la información de sincronía. Saltar entre NLE puede provocar pérdida de datos —por ejemplo, códigos de tiempo o nombres de archivo—, de modo que apartarse del *pipeline* unificado añadió pasos extra para volver a alinear contenidos.

Otro problema habitual fue la ausencia de código de tiempo en algunas grabaciones de ambiente realizadas con grabadoras de campo independientes. Cuando una cámara o una grabadora opera sin reloj sincronizado, se generan discrepancias que luego deben corregirse manualmente. Por eso es fundamental que todo el material, sin excepción, incluya código de tiempo o, al menos, una referencia clara de sincronía, garantizando una integración fluida de todas las fuentes de audio y video en la postproducción.

La figura 1 resume el flujo de trabajo completo de postproducción que se debió aplicar en este proyecto, desde la fase de rodaje hasta la entrega final del documental.

---

6 Huber, Runstein y Holmes, *Audio Post Production...*

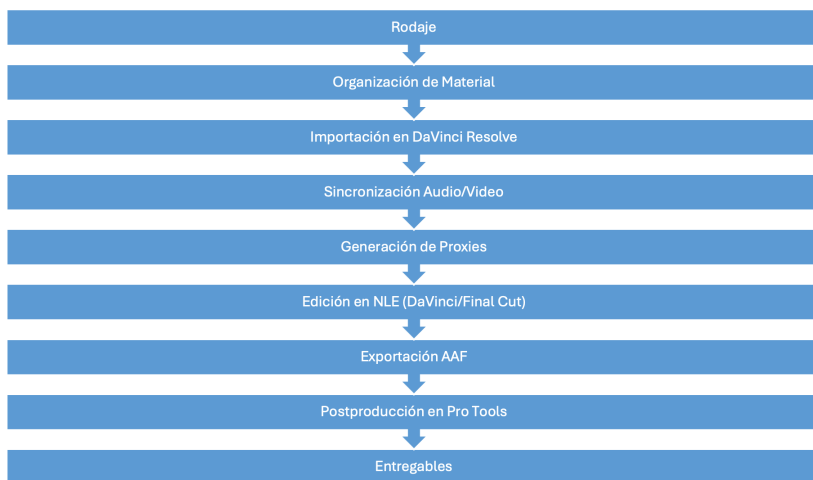


Figura 1. Diagrama de flujo del proceso de postproducción. Fuente: Elaboración propia.

## Problemas y soluciones

Durante la postproducción se identificaron casos en los que, a pesar de contar con código de tiempo, existían desfases sistemáticos entre audio y video. Para corregirlos fue necesario modificar manualmente los valores de *timecode* en DaVinci Resolve. Con el fin de reducir errores y tareas repetitivas, se desarrolló una calculadora de código de tiempo que automatiza la suma y resta de fotogramas.

Cuando el desfase se identificó mediante la forma de onda, se extrajo el diferencial entre códigos de tiempo y se reutilizó en todas las grabaciones realizadas bajo las mismas condiciones. Ello agilizó el proceso: bastó confirmar que el ajuste funcionaba para cada clip, sin repetir la sincronización desde cero, logrando una alineación uniforme en todo el material.

La calculadora de *timecodes* está disponible como aplicación web interactiva (<http://tccalc.vicherrera.net/>) y su código fuente se publica en GitHub (<https://github.com/vhhr7/tccalculator>), de modo que cualquier editor puede corregir desajustes de forma rápida y precisa.

Asimismo, en situaciones donde ni el código de tiempo ni la forma de onda permitían una sincronización fiable, se implementó una herramienta de transcripción automática basada en el modelo Whisper, desarrollada en Python con interfaz en Gradio, su código fuente se encuentra en <https://github.com/vhhr7/transcript>. Esta solución permitió localizar eventos sonoros específicos mediante texto con marcas temporales, facilitando la sincronización en fragmentos carentes de referencias claras.

Radford *et al.* describen que «los modelos de reconocimiento automático del habla entrenados mediante supervisión débil a gran escala presentan una alta robustez frente a ruido, variabilidad acústica y condiciones de grabación no controladas»<sup>7</sup>. Esto permite su aplicación efectiva en contextos documentales reales donde el material sonoro no responde a condiciones de estudio.

El flujo es sencillo: la aplicación carga varios archivos de audio, los transcribe con Whisper señalando inicio y fin de cada segmento en formato HH\MM\SS\FF y exporta un CSV que permite buscar palabras clave. Así, los editores localizan de inmediato los pasajes relevantes cuando otros métodos fallan, simplificando tanto la búsqueda como la sincronización posterior.

En conjunto, estas herramientas reducen tareas manuales repetitivas y elevan la precisión del resultado final, haciendo el flujo de postproducción documental mucho más eficiente.

25

## Consideraciones éticas sobre el uso de inteligencia artificial

El uso de herramientas basadas en inteligencia artificial en el desarrollo de este artículo se empleó como un recurso editorial de apoyo, orientado a mejorar la coherencia, claridad y fluidez de la redacción, así como a la organización formal del texto. Dichas herramientas no sustituyeron el tra-

---

<sup>7</sup> Alec Radford, *et al.*, «Robust Speech Recognition via Large-Scale Weak Supervision», *Arxiv* (2022), <https://arxiv.org/abs/2212.04356>

bajo intelectual del autor ni intervinieron en la generación del contenido analítico, los argumentos teóricos o las conclusiones presentadas.

El artículo se fundamenta en la resolución de un problema real de sincronización audiovisual surgido durante la postproducción de una serie documental actualmente en proceso de estreno, contexto que habilita al autor a reflexionar críticamente sobre el caso sin comprometer información sensible del proyecto. El análisis, la interpretación de los resultados y la toma de decisiones técnicas se basaron íntegramente en la práctica profesional, la experiencia directa del autor y el uso de tecnologías disponibles que hicieron posible avanzar y resolver las dificultades descritas.

La utilización de inteligencia artificial se declara de manera explícita con el fin de garantizar la transparencia metodológica y cumplir con los principios de honestidad académica exigidos en la producción científica contemporánea.

## 26

# Resultados/conclusiones

## Desafíos y resultados

La sincronización audiovisual en documentales presenta desafíos específicos derivados de la diversidad técnica y la naturaleza impredecible del rodaje. Este estudio demuestra que la implementación de flujos de trabajo estandarizados, el uso consistente de código de tiempo y la incorporación de herramientas de apoyo basadas en inteligencia artificial pueden reducir significativamente los retrabajos en postproducción.

Para investigaciones futuras se recomienda profundizar en soluciones automatizadas que integren análisis audiovisual y aprendizaje automático, así como evaluar su impacto en entornos de producción de mayor escala.

Para proyectos venideros se recomienda un flujo de trabajo estandarizado que incluya:

- Uso consistente de dispositivos con código de tiempo en todas las cámaras y grabadoras de audio, garantizando una referencia temporal común desde el rodaje.
- Verificación diaria de la precisión de la sincronía al inicio de cada jornada (señal de audio compartida y comparación de *timecodes*) para detectar desviaciones cuanto antes.
- Organización rigurosa de archivos y metadatos, etiquetando material, conservando fecha y hora, y documentando cualquier ajuste realizado, de modo que el equipo pueda rastrear decisiones y facilitar resincronizaciones futuras.

Aplicar estas medidas minimiza los problemas de sincronización, optimiza el tiempo de edición y, al establecer un protocolo único para todo el equipo, evita incompatibilidades de flujo y preserva la información de sincronía a lo largo de todo el pipeline.

27

## Conclusión

La sincronización de audio en documentales supone desafíos particulares, derivados de la producción heterogénea y espontánea de este género. Este estudio demuestra que la ausencia de un protocolo unificado genera retrabajos significativos en postproducción; en cambio, un flujo estandarizado y buenas prácticas —uso constante de código de tiempo y organización meticulosa del material— reducen de forma sustancial los tiempos de edición y mejoran la eficiencia global. Para investigaciones futuras se recomienda explorar soluciones automatizadas más avanzadas, basadas en inteligencia artificial, que identifiquen y alineen audio y video incluso en entornos de grabación adversos. Así, la industria documental podría disponer de herramientas más robustas y veloces para afrontar los retos de sincronización.

Cómo citar este artículo:

Herrera, Víctor. «Cuando el timecode falla: inteligencia artificial y estrategias alternativas para la sincronización en documentales». *Fuera de Campo* 8, n.º 2 (2024): 08-28.

## Bibliografía

28

Rapp, Henri. «How to Sync Audio for Modern Video Production Workflows». No Film School, 14 de diciembre de 2020. <https://nofilmschool.com/how-sync-audio-modern-video-production-workflows> (acceso el 2 de abril de 2025).

Herrera, Víctor. Timecode Calculator (aplicación web). 2023. <https://tccalc.vicherrera.net> (acceso el 2 de abril de 2025).

Herrera, Víctor. Timecode Calculator-Código Fuente (repositorio en GitHub). 2023. <https://github.com/vhhr7/tccalculator> (acceso el 2 de abril de 2025).

Herrera, Víctor. Herramienta de transcripción automática-Código Fuente (repositorio en GitHub). 2025. <https://github.com/vhhr7/transcript> (acceso el 2 de abril de 2025).

Society of Motion Picture and Television Engineers. «SMPTE ST 12-1:2014-Time and Control Code. White Plains». SMPTE, 2014.

Eg, Ragnhild y Dawn M. Behne. «Perceived Synchrony for Realistic and Dynamic Audiovisual Events». *Frontiers in Psychology* 6, n.º 736 (2015): 1-12. <https://doi.org/10.3389/fpsyg.2015.00736>.

Huber, David Miles, Robert E. Runstein y Richard Holmes. *Audio Post Production for Television and Film: An Introduction to Technology and Techniques*. 3.<sup>a</sup> ed. Oxford: Focal Press, 2004.

Radford, Alec, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, y otros. «Robust Speech Recognition via Large-Scale Weak Supervision». *Arxiv* (2022). <https://arxiv.org/abs/2212.04356>.