# La Guía Definitiva de Prompting para Veo 3.1

Si una imagen vale más que mil palabras, un video vale un millón. Para los creadores, el video generativo promete dar vida a cualquier historia. Sin embargo, la realidad a menudo ha sido un ciclo frustrante de "probar y rezar", con poco control sobre la consistencia de los personajes, la calidad cinematográfica o la coherencia narrativa.

Esta guía es tu marco de trabajo para *dirigir* **Veo 3.1**, nuestro último modelo que marca el cambio de la simple generación al control creativo. Veo 3.1 se basa en Veo 3, con una adherencia a los prompts más fuerte y una calidad audiovisual mejorada, especialmente al convertir imágenes estáticas en videos.

**Qué aprenderás en esta guía:**

* La gama completa de capacidades de Veo 3.1.
* Una fórmula para dirigir escenas con personajes y estilos consistentes.
* Cómo dirigir el video y el sonido utilizando técnicas cinematográficas profesionales.
* Cómo ejecutar ideas complejas combinando Veo con modelos de imagen de Gemini en flujos de trabajo avanzados.

## 1. Capacidades del Modelo Veo 3.1

Primero, es esencial entender el conjunto completo de funciones del modelo. Veo 3.1 añade audio a las capacidades existentes para ayudarte a crear la escena perfecta.

### Funciones Principales de Generación

* **Video de alta fidelidad:** Genera video a resolución de 720p o 1080p.
* **Relación de aspecto:** 16:9 (horizontal) o 9:16 (vertical).
* **Duración variable:** Crea clips de 4, 6 u 8 segundos.
* **Audio y diálogo enriquecidos:** Veo 3.1 destaca en la generación de sonido realista y sincronizado, desde conversaciones de varias personas hasta efectos de sonido precisos, todo guiado por el prompt.
* **Comprensión de escenas complejas:** El modelo tiene una comprensión más profunda de la estructura narrativa y los estilos cinematográficos, lo que le permite representar mejor las interacciones de los personajes y seguir las indicaciones de la historia.

### Controles Creativos Avanzados

* **Imagen-a-Video (Mejorado):** Anima una imagen fuente con mayor adherencia al prompt y una calidad audiovisual mejorada.
* **Elementos Consistentes ("Ingredientes para video"):** Proporciona imágenes de referencia de una escena, personaje, objeto o estilo para mantener una estética coherente en múltiples tomas. Esta función ahora incluye la generación de audio.
* **Transiciones fluidas ("Primer y último fotograma"):** Genera una transición de video natural entre una imagen de inicio y una imagen de final proporcionadas, con audio completo.
* **Añadir/Eliminar Objeto:** Introduce nuevos objetos o elimina los existentes de un video generado. Veo preserva la composición original de la escena.
* **Marca de agua digital:** Todos los videos generados se marcan con SynthID para indicar que el contenido es generado por IA.

**Nota importante:** La función Añadir/eliminar objeto actualmente utiliza el modelo Veo 2 y, por lo tanto, no genera audio.

## 2. Una Fórmula para Prompts Efectivos

Un prompt estructurado produce resultados consistentes y de alta calidad. Considera esta fórmula de cinco partes para un control óptimo:

$$[Cinematografía] + [Sujeto] + [Acción] + [Contexto] + [Estilo \text{ y } Ambiente]$$

* **Cinematografía:** Define el trabajo de cámara y la composición de la toma.
* **Sujeto:** Identifica al personaje principal o punto focal.
* **Acción:** Describe lo que está haciendo el sujeto.
* **Contexto:** Detalla el entorno y los elementos de fondo.
* **Estilo y Ambiente:** Especifica la estética general, el estado de ánimo y la iluminación.

Ejemplo de prompt:

Medium shot, un trabajador de oficina cansado, frotándose las sienes con agotamiento, frente a una voluminosa computadora de 1980 en una oficina desordenada a altas horas de la noche. La escena está iluminada por las duras luces fluorescentes del techo y el brillo verde del monitor monocromático. Estética retro, filmado como en película de color de 1980, ligeramente granulado.

## 3. Técnicas Esenciales de Prompting

Dominar estas técnicas básicas te dará un control granular sobre cada aspecto de tu generación.

### El Lenguaje de la Cinematografía

El elemento [Cinematografía] de tu prompt es la herramienta más poderosa para transmitir tono y emoción.

* **Movimiento de cámara:** Dolly shot, tracking shot (toma de seguimiento), crane shot (toma de grúa), aerial view (vista aérea), slow pan (paneo lento), POV shot (toma subjetiva).
* **Composición:** Wide shot (plano general), close-up (primer plano), extreme close-up, low angle (ángulo contrapicado), two-shot (plano de dos).
* **Lente y enfoque:** Shallow depth of field (poca profundidad de campo), wide-angle lens (lente gran angular), soft focus (enfoque suave), macro lens (lente macro), deep focus (enfoque profundo).

Ejemplo (Toma de grúa):

Toma de grúa (crane shot) comenzando bajo sobre un excursionista solitario y ascendiendo muy por encima, revelando que está de pie al borde de un cañón colosal lleno de niebla al amanecer, estilo de fantasía épica, impresionante, luz suave de la mañana.

Ejemplo (Poca profundidad de campo):

Primer plano (close-up) con muy poca profundidad de campo (shallow depth of field), el rostro de una mujer joven, mirando por la ventana de un autobús las luces de la ciudad que pasan, con su reflejo apenas visible en el cristal, dentro de un autobús por la noche durante una tormenta, ambiente melancólico con tonos azules fríos, melancólico, cinematográfico.

### Dirigiendo el Escenario Sonoro

Veo 3.1 puede generar una banda sonora completa basada en tus instrucciones de texto.

* **Diálogo:** Usa comillas para diálogos específicos (ej. Una mujer dice: "Tenemos que irnos ahora").
* **Efectos de sonido (SFX):** Describe los sonidos con claridad (ej. SFX: truenos retumban en la distancia).
* **Ruido ambiental:** Define el paisaje sonoro de fondo (ej. Ruido ambiental: el zumbido silencioso del puente de una nave estelar).

Divide el audio en tres niveles:

- fx: sonidos físicos o mecánicos

- music: partitura o motivo emocional

- ambient: paisaje sonoro contextual

### Dominando los Prompts Negativos

Para refinar tu resultado, describe lo que deseas excluir. Especifica claramente lo que *no* quieres ver. Por ejemplo, en lugar de "sin estructuras hechas por el hombre", usa un paisaje desolado sin edificios ni carreteras.

### Mejora de Prompts con Gemini

Si necesitas añadir más detalles, usa un modelo como Gemini para analizar y enriquecer un prompt simple con un lenguaje más descriptivo y cinematográfico.

## 4. Flujos de Trabajo Creativos Avanzados

Si bien un solo prompt detallado es poderoso, un flujo de trabajo de varios pasos ofrece un control inigualable. Los siguientes flujos demuestran cómo combinar las nuevas capacidades de Veo 3.1 con un modelo de generación de imágenes (como Gemini Flash Image) para ejecutar visiones complejas.

### Flujo 1: La Transición Dinámica ("Primer y último fotograma")

Esta técnica te permite crear un movimiento de cámara o transformación específica y controlada entre dos puntos de vista distintos.

1. **Crea el fotograma inicial:** Usa Gemini (o tu herramienta de imagen preferida) para generar tu toma inicial.
   * *Prompt de Gemini (ejemplo):* Medium shot de una estrella de pop femenina cantando apasionadamente en un micrófono vintage. Está en un escenario oscuro, iluminada por un único y dramático foco desde el frente. Tiene los ojos cerrados, capturando un momento emocional. Fotorrealista, cinematográfico.
2. **Crea el fotograma final:** Genera una segunda imagen complementaria.
   * *Prompt de Gemini (ejemplo):* Toma subjetiva (POV) desde detrás de la cantante en el escenario, mirando a una gran multitud que vitorea. Las luces del escenario son brillantes, creando destellos de lente. Se puede ver la parte posterior de la cabeza y los hombros de la cantante en primer plano. La audiencia es un mar de luces y siluetas. Atmósfera energética.
3. **Anima con Veo:** Introduce ambas imágenes en Veo usando la función Primer y último fotograma. En tu prompt, describe la transición y el audio que deseas.
   * *Prompt de Veo 3.1:* La cámara realiza una toma de arco suave de 180 grados, comenzando con la vista frontal de la cantante y girando a su alrededor para terminar sin problemas en la toma subjetiva desde detrás de ella en el escenario. La cantante canta "cuando me miras a los ojos, puedo ver un millón de estrellas".

### Flujo 2: Construyendo una Escena de Diálogo ("Ingredientes para video")

Este flujo es ideal para crear una escena de múltiples tomas con personajes consistentes en una conversación, aprovechando la habilidad de Veo 3.1 para crear diálogos.

1. **Genera tus "ingredientes":** Crea imágenes de referencia con Gemini para tus personajes y el escenario (ej. imagen\_detective, imagen\_mujer, imagen\_oficina).
2. **Compón la escena (Toma 1):** Usa la función Ingredientes para video con las imágenes de referencia relevantes.
   * *Prompt:* Usando las imágenes proporcionadas para el detective, la mujer y el entorno de la oficina, crea un medium shot del detective detrás de su escritorio. Él mira a la mujer y dice con voz cansada: "De todas las oficinas en esta ciudad, tenías que entrar en la mía".
3. **Compón la escena (Toma 2):**
   * *Prompt:* Usando las imágenes proporcionadas para el detective, la mujer y el entorno de la oficina, crea una toma centrada en la mujer. Una ligera y misteriosa sonrisa se dibuja en sus labios mientras responde: "Me dijeron que eras el mejor".

### Flujo 3: Prompting con Marcas de Tiempo (Timestamping)

Este flujo te permite dirigir una secuencia completa de varias tomas con una sincronización cinematográfica precisa, todo dentro de una sola generación. Al asignar acciones a segmentos de tiempo, puedes crear eficientemente una escena completa.

**Ejemplo de prompt:**

[00:00-00:02] Medium shot desde detrás de una joven exploradora con un bolso de cuero y cabello castaño desordenado en una cola de caballo, mientras aparta una gran liana de la jungla para revelar un camino oculto.

[00:02-00:04] Plano inverso del rostro pecoso de la exploradora, su expresión llena de asombro mientras contempla unas antiguas ruinas cubiertas de musgo al fondo. SFX: El susurro de hojas densas, llamadas de aves exóticas distantes.

[00:04-00:06] Tracking shot siguiendo a la exploradora mientras entra en el claro y pasa la mano sobre los intrincados grabados de un muro de piedra desmoronado. Emoción: Maravilla y reverencia.

[00:06-00:08] Toma de grúa (crane shot) amplia y en ángulo picado, revelando a la exploradora solitaria, pequeña en el centro del vasto complejo del templo olvidado, semi-tragado por la jungla. SFX: Una suave partitura orquestal comienza a sonar y crecer.

Control Temporal y Loop Cinemático

Incorpora bloques de tiempo (“[00:00–00:03]”, etc.) para definir acciones, emociones y transiciones de luz/sonido.

Añade "loop\_hint: true" cuando desees un final que regrese suavemente al inicio (ideal para reels o fondos).

Reglas Físicas y Visuales

physics\_rules: cómo interactúan polvo, vapor, luz o partículas con el entorno.

visual\_rules: evita errores de renderizado (cartoonish, distorsiones, desenfoques).

**Flujos de trabajo creativos avanzados**

Incorporar “Loop y Arco Sonoro”: Usa timestamp + audio evolutivo para construir crescendos visuales o loops atmosféricos.